

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
11 October 2001 (11.10.2001)

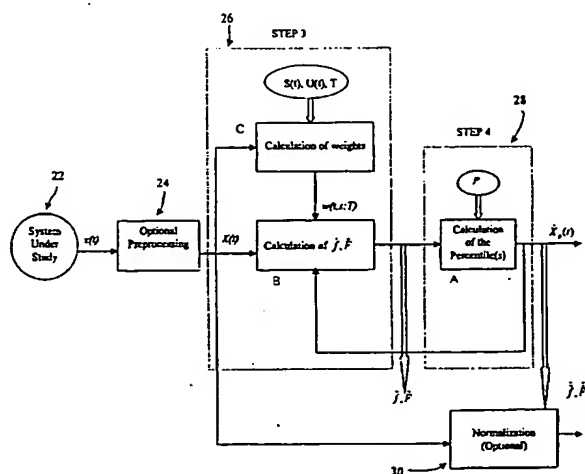
PCT

(10) International Publication Number  
**WO 01/75660 A1**

- (51) International Patent Classification<sup>7</sup>: **G06F 17/13**, 17/14
- (74) Agent: **LUEBBERING, Thomas, B.**; Hovey, Williams, Timmons & Collins, Suite 400, 2405 Grand Boulevard, Kansas City, MO 64108 (US).
- (21) International Application Number: **PCT/US01/10677**
- (22) International Filing Date: **3 April 2001 (03.04.2001)**
- (25) Filing Language: **English**
- (26) Publication Language: **English**
- (30) Priority Data:  
60/194,130 3 April 2000 (03.04.2000) US  
Not furnished 3 April 2001 (03.04.2001) US
- (71) Applicant: **FLINT HILLS SCIENTIFIC, L.L.C.**  
[US/US]; Suite A, 5020 W. 15th Street, Lawrence, KS 66049 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- (72) Inventors: **NIKITIN, Alexei, V.**; 2124 Vermont Avenue, Lawrence, KS 66046 (US). **FREI, Mark, G.**; 2513 Via Linda Drive, Lawrence, KS 66047 (US). **BHAVARAJU, Naresh, C.**; 209 V V Camden Manor, Lawrence, KS 66049 (US). **OSORIO, Ivan**; 4005 W. 124th Street, Leawood, KS 66209 (US).
- Published:  
— with international search report  
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

[Continued on next page]

(54) Title: **METHOD, COMPUTER PROGRAM, AND SYSTEM FOR AUTOMATED REAL-TIME SIGNAL ANALYSIS FOR DETECTION, QUANTIFICATION, AND PREDICTION OF SIGNAL CHANGES**



(57) Abstract: A method and system for real-time signal analysis providing characterization of temporally-evolving densities and distributions (26) of signal features of arbitrary-type signals (22) in a moving time window by tracking output of order statistic filters (28) (also known as percentile, quantile, or rank-order filters). Given a raw input signal of arbitrary type, origin, or scale, the present invention enables automated quantification and detection of changes in the distribution of any set of quantifiable features of that signal as they occur in time. Furthermore, the present invention's ability to rapidly and accurately detect changes in certain features of an input signal can also enable prediction in cases where the detected changes associated with an increased likelihood of future signal changes.

WO 01/75660 A1



---

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

METHOD, COMPUTER PROGRAM, AND SYSTEM FOR AUTOMATED REAL-TIME SIGNAL ANALYSIS FOR  
DETECTION, QUANTIFICATION, AND PREDICTION OF SIGNAL CHANGES

5

10

## 15 BACKGROUND OF THE INVENTION

## 1. FIELD OF THE INVENTION

The present invention relates generally to methods and systems for automated signal analysis providing rapid and accurate detection, prediction, or quantification of changes in one or more signal features, characteristics, or properties as they occur. More particularly, the present invention relates to a method or system for automated real-time signal analysis providing characterization of temporally-evolving densities and distributions of signal features of arbitrary-type signals in a moving time window by tracking output of order statistic filters (also known as percentile, quantile, or rank-order filters).

25

## 2. DESCRIPTION OF THE PRIOR ART

It is often desirable to detect and quantify feature changes in an evolving signal, and there have been numerous attempts to develop automated signal analysis means operable to do so. One well-known approach, for example, is based upon analysis of the signal's mean value, which is typically a well known, well understood, and easily computed property. Other well-known techniques look for changes in signal variance or standard deviation over time.

30

Unfortunately, these commonly used approaches have significant drawbacks, including lack of robustness in the presence of signal outliers. Furthermore, in all but

a few ideal cases, monitoring these individual parameters does not enable detection of all types of changes in feature distribution. This is because the mean and standard deviation rarely completely describe the signal distribution. Another problem that plagues many existing analysis techniques is that they are unable to  
5 deal adequately with real-world problems in which the analyzed signal is often highly complex, non-stationary, non-linear, and/or stochastic.

Another well-known approach, one more suited to practical situations than the above-mentioned methods, uses order statistics (e.g., the median or other percentile or quantile values). Order statistics are advantageous because they are directly  
10 related to the underlying distribution and are robust in the presence of outliers. For example, a method of signal analysis that enables the detection of state changes in the brain through automated analysis of recorded signal changes is disclosed in U.S. Patent No. 5,995,868. This method addresses the problem of robustness in the presence of outliers through novel use of order-statistic filtering. Additionally, given  
15 information from a moving time window of a certain time scale, referred to as the "foreground", this method provides for real-time comparison thereof with a reference obtained from past data derived, e.g., from a longer time scale window, referred to as the "background." This approach thereby addresses some of the normalization problems associated with complex, non-stationary signals.

20 Although the prior invention disclosed in U.S. Patent No. 5,995,868 has successfully addressed many of the above-mentioned limitations, including normalization problems associated with complex non-stationary signals, it is lacking in breadth of scope. Detection of changes, for example, is limited to a particular order statistic of the signal. Additionally, the order statistic filter employed to detect  
25 signal changes requires large amounts of processing ability, memory, and power when used on digital signals for which sorting procedures are performed at each point in time. Furthermore, the method does not enable full analog implementation.

Most work on order statistic filters, such as median filters, and their implementation is in the areas of digital signal and image processing, which, as  
30 mentioned, requires large amounts of processing ability, memory, and power not practical or cost-effective for some applications. Work on analog median filters is limited to situations where the input is provided as parallel lines of data, and a program or circuit that implements the filter outputs a value that is equal to the median of the data on different input lines. Work on analog median filtering for

continuous-time signals is not extensive, and no realizable implementations exist able to track a percentile (e.g., median) of a continuous-time signal. One reason for this is that the operation of finding the rank or order is non-linear; making modeling the procedure using an ordinary differential equation so complicated that it has not yet been addressed.

Due to the above-described and other problems, a need exists for a more general, powerful, and broad method for automated analysis of signals of any degree of complexity and type.

## 10 SUMMARY OF THE INVENTION

The present invention solves the above-described and other problems to provide a distinct advance in the art of automated signal analysis. More specifically, the present invention comprises a method and system for real-time signal analysis providing characterization of temporally-evolving densities and distributions of signal features of arbitrary-type signals in a moving time window by tracking output of order statistic filters (also called percentile, quantile, or rank-order filters).

The present invention is operable to analyze input signals of arbitrary type, origin and scale, including, for example, continuous-time or discrete-time, analog or digital, scalar or multi-dimensional, deterministic or stochastic (i.e., containing a random component), stationary (i.e., time invariant) or non-stationary (i.e., time varying), linear or nonlinear. Thus, the present invention has broad applicability to analysis of many different types of complex signals and sequences of data, including but not limited to biological signals such as those produced by brain, heart, or muscle activity; physical signals such as seismic, oceanographic, or meteorological; financial signals such as prices of various financial instruments; communication signals such as recorded speech or video or network traffic signals; mechanical signals such as jet engine vibration; target tracking and recognition; signals describing population dynamics, ecosystems or bio-systems; signals derived from manufacturing or other queuing systems; chemical signals such as spectroscopic signals; and sequences of data such as word lists, documents, or gene sequences. Furthermore, the present invention is applicable to any set of signal features so long as they are quantifiable, thereby allowing for a high degree of system adaptability and selectivity.

Thus, the present invention enables automated detection and quantification of changes in the distribution of any set of quantifiable features of a raw input signal as they occur in time. The input signal, denoted as  $\{x(t)\}$ , can be any data parameterized by a real-valued variable,  $t$ , which will be interpreted as a time variable. The input signal may be optionally preprocessed in order to produce a new signal, the feature signal, denoted as  $\{X(t)\}$ .  $\{X(t)\}$  quantifies a set of features of the input signal that the system will use in detecting and quantifying changes. For a fixed  $t$ ,  $X(t)$  is called the signal feature vector at time  $t$ . The feature vector has as many components as there are signal features. While potentially of substantial value, this preprocessing step is optional in the sense that the raw input signal itself may be used as the feature vector (i.e.,  $X(t) = x(t)$ ), in which case the invention proceeds to detect changes in the distribution of the raw input signal as it evolves in time. The desirability of preprocessing will depend upon the nature of the raw input signal and the nature of the features of interest.

The present invention also introduces a useful new object called the time-weighted feature density of a signal,  $\{f(t,X)\}$ , which can be computed from the feature signal at each point in time. This object allows access to estimates of the full time-dependent density and cumulative distribution function of varying signal features with any desired degree of accuracy, but confines these estimates to any desired time-scale through the use of time-weighting (time localization of feature density). This time-weighted feature density describes the raw input signal features measured in moving windows of time specified by the time-weight function, which allows a user to apply different significances to portions of available information (e.g., to consider recent information as more relevant than older information; or to weight information according to its reliability, etc.).

Moreover, the present invention allows for rapidly obtaining these estimates in a computationally efficient manner that can be implemented in digital or analog form, and a method for detecting, quantifying, and comparing changes of arbitrary type in the density/distribution of the feature vector as it changes. The significance of this increase in computational efficiency, along with analog implementability, becomes especially clear when considering medical device applications where, for example, the present invention enables currently used externally-worn devices that

require daily battery recharging to become fully implantable devices with an operational lifetime of several years, thereby improving safety and convenience.

In operation, a raw time-varying input signal of arbitrary type, origin, and scale is received for analysis. Optionally, depending upon the nature of the raw input signal and the nature of the features of interest, pre-processing occurs to produce a feature signal more amenable to further analysis. Next, time-weighted density or distribution functions are determined for both a foreground or current time window portion of the signal and a background portion of the signal or reference signal (which also may be evolving with time, but potentially on a different timescale) in order to emphasize, as desired, certain data.

Percentile values for both foreground and background signals are then accurately estimated and compared so as to detect and quantify feature changes on any timescale and to any desired degree of precision as the raw input signal evolves in time. Density and distribution approximations may also be compared. As noted above, the state of the existing art requires that the data be laboriously sorted in order to determine these percentile values. In the present invention, however, percentile values are accurately estimated without sorting or stacking, thereby increasing processing speed and efficiency while reducing computation, memory, and power needs. Thus, the present invention is able to perform in a highly computationally efficient manner that can be implemented in a low power consumption apparatus consisting of an analog system, a digital processor, or a hybrid combination thereof, thereby providing tremendous system power savings.

The present invention is also operable to facilitate real-time signal normalization with respect to the density/distribution approximations, which is useful in processing and analysis of series of different orders. This is particularly useful where the features or characteristics of interest are invariant to a monotonic transformation of the signal's amplitude.

It will also be appreciated that the present invention's ability to rapidly and accurately detect changes in certain features of the input signal can enable prediction in cases where the changes it detects are associated with an increased likelihood of future signal changes. For example, when applied to seismic signals, the method can enable prediction of an earthquake or volcanic eruption; when applied to meteorological signals, the method can enable prediction of severe weather; when applied to financial data, the method can enable prediction of an

impending price change in a stock; when applied to brain waves or heart signals, the method can enable prediction of an epileptic seizure or ventricular fibrillation; and when applied to brain wave or electromyographic signals, it can enable prediction of movement of a body part.

- 5           These and other novel features of the present invention are described in more detail below in the section titled DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT.

#### BRIEF DESCRIPTION OF THE DRAWING FIGURES

- 10           The present invention is described in detail below with reference to the attached drawing figures, wherein:

FIG. 1 is a block diagram illustrating a first portion of steps involved in performing a preferred embodiment of the present invention;

- 15           FIG. 2 is a block diagram illustrating a second portion of steps involved in performing a preferred embodiment of the present invention;

FIG. 3 is a block diagram illustrating a third portion of steps involved in performing a preferred embodiment of the present invention;

FIG. 4 is a graph of an exemplary raw input signal,  $x(t)$ , as might be received for analysis by a preferred embodiment of the present invention;

- 20           FIG. 5 is a graph of a feature signal,  $X(t)$ , resulting from preprocessing the raw input signal shown in FIG. 4;

FIG. 6 is a graph showing the calculated 0.25, 0.50, and 0.75 percentiles of the feature signal shown in FIG. 5;

- 25           FIG. 7 is a graph showing the true feature density,  $f(t,w)$  of the feature signal shown in FIG. 5 calculated at times  $t_1$  and  $t_2$ ;

FIG. 8 is a graph showing the true feature distribution,  $F(t,w)$ , of the feature signal shown in FIG. 5 calculated at times  $t_1$  and  $t_2$ ;

FIG. 9 shows a graph of an evolving first approximation of the time-weighted feature density of the feature signal shown in FIG. 5, calculated at times  $t_1$  and  $t_2$ ;

- 30           FIG. 10 shows a graph of an evolving first approximation of the time-weighted distribution function of the feature signal shown in FIG. 5, calculated at times  $t_1$  and  $t_2$ ;

FIG. 11 is a graph showing calculated percentile tracking filter outputs for 0.25, 0.50, and 0.75 percentiles of the feature signal shown in FIG. 5;



FIG. 12 shows a graph of an evolving second approximation of the time-weighted feature density of the feature signal shown in FIG. 5, calculated at times  $t_1$  and  $t_2$ ;

FIG. 13 shows a graph of an evolving second approximation of the time-weighted distribution function of the feature signal shown in FIG. 5, calculated at times  $t_1$  and  $t_2$ ;

FIG. 14 shows a graph of a  $\Lambda(t)$  measured from the feature signal shown in FIG. 5;

FIG. 15 is a block diagram of a preferred embodiment of an analog implementation of a percentile tracking filter component of the present invention;

FIG. 16 is a detailed circuit schematic of the percentile tracking filter component shown in FIG. 15;

FIG. 17A shows an exemplary feature signal that for analysis by the present invention;

FIG. 17B shows an output of the detailed circuit schematic shown in FIG. 16 and a true median output associated with the feature signal of FIG. 17A;

FIG. 18 is a block diagram of a preferred embodiment of an analog implementation of a Lambda estimator component of the present invention; and

FIG. 19 is a detailed circuit schematic of the Lambda estimator component shown in FIG. 18.

#### DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

The present invention comprises a method and system for real-time signal analysis providing characterization of temporally-evolving densities and distributions of signal features of arbitrary-type signals in a moving time window by tracking output of order statistic (e.g., percentile, quantile, rank-order) filters. More specifically, given a raw input signal of arbitrary type, origin, and scale, the present invention enables automated quantification and detection of changes in the distribution of any set of quantifiable features of that signal as they occur in time. Furthermore, the present invention's ability to rapidly and accurately detect changes in certain features of an input signal can also enable prediction in cases when the detected changes are associated with an increased likelihood of future signal changes.

## Method

### Step 1: Receive Raw Input Signal

Referring to FIG. 1, the raw input signal is received from a system under study

5 22. This signal, denoted as  $\{x(t)\}$ , can be any data parameterized by a real-valued variable,  $t$ , which will be interpreted as a time variable. As noted, the raw input signals may be of arbitrary type, including, for example, continuous-time or discrete-time, analog or digital, scalar or multi-dimensional, deterministic or stochastic (i.e., containing a random component), stationary (i.e., time invariant) or non-stationary

10 (i.e., time varying), and linear or nonlinear. The raw input signals may also be of arbitrary origin, including, for example, biological signals such as those produced by brain, heart, or muscle activity; financial signals such as prices of various financial instruments; physical signals such as seismic, oceanographic, and meteorological; communication signals such as recorded speech or video or network traffic signals;

15 mechanical signals such as jet engine vibration; chemical signals such as those obtained in spectroscopy; and sequences of data such as word lists or gene sequences.

### Step 2: (Optional) Preprocess Raw Input Signal to Derive Feature Signal

20 The raw input signal may be optionally pre-processed, as shown in box 24, in order to produce a new signal, the feature signal, denoted as  $\{X(t)\}$ . One skilled in the art will appreciate that there exist a nearly endless set of transformations that can be applied to the raw input signal,  $x(t)$ , to quantify various features, characteristics, or properties of the signal as it evolves. Common examples include derivatives of any

25 order; integrals or any order; various moments and related properties such as variance, skewness, kurtosis; wavespeed and related measures such as inter-zero-crossing intervals and inter-peak intervals; signal power in a time window and/or in a particular frequency band; measures derived from Fourier analysis such as those involving signal phase or power spectral density; measures from nonlinear dynamics

30 such as correlation dimension, fractal dimension, magnitude of Lyapunov exponents; phase delay embeddings; and measures of rhythmicity, wave shape, or amplitude.

The feature signal,  $\{X(t)\}$ , quantifies a set of features of the input signal that the system will use in detecting changes. For a fixed  $t$ ,  $X(t)$  is called the signal

feature vector at time  $t$ . The feature vector has as many components as there are signal features. While potentially of substantial value, this pre-processing step is optional in the sense that the raw input signal itself may be used as the feature vector (i.e.,  $X(t) = x(t)$ ), in which case the invention proceeds to detect changes in the distribution of the raw signal as it evolves in time. The desirability of preprocessing will depend upon the nature of the raw input signal and the nature of the features of interest.

Step 3: Determine Time-Weighted Distribution and Density Functions of the Feature Signal

A time-weighted feature density (TWFD) of a signal,  $\{f(t, X)\}$ , is computed from the raw input signal or feature signal at each point in time, as shown in box 26. The TWFD allows access to estimates of the full time-dependent density and cumulative distribution functions of varying signal features with any desired degree of accuracy, but confines these estimates to any desired time-scale through the use of time-weighting (time localization of feature density). Time-weighting allows the user to apply different significance to portions of available information (e.g., to consider more recent information as more relevant than older information; or to weigh information according to its reliability, etc.).

Thus, the TWFD describes the raw input signal features measured in moving windows of time specified by a time-weight function, thereby allowing for detection, quantification, and comparison of changes of arbitrary type in the density/distribution of the feature vector as it evolves.

The instantaneous feature density of a signal,  $X$ , at time  $t$  is defined as

$$f(t, D) = \delta[D - X(t)], \quad (1)$$

where  $\delta(x)$  is the Dirac  $\delta$ -function, and  $D$  is a variable signal level. This density is just a  $\delta$ -function at the signal's present value.

The concept of instantaneous feature density can then be extended to a time-weighted window. The TWFD of the feature signal  $X(t)$  in a time window  $w(t, s)$  is defined as

$$f(t, D) = \int_{-\infty}^{\infty} w(t, s) \delta[D - X(s)] ds, \quad (2)$$

where  $\delta(x)$  is the Dirac  $\delta$ -function,  $D$  is the amplitude, and the weighting function,  $w(t,s)$ , often called a "time-weight," is any function such that

$$\int_{-\infty}^{\infty} w(t,s) ds = 1. \quad (3)$$

Typically,  $w(t,s) \geq 0$  for all  $t,s$ . Also, in most practical applications,  $w$  will be "causal" meaning that  $w(t,s) = 0$  for all  $s > t$ . It should be noted that  $w(t,s)$  may attain its dependence upon  $t$  or  $s$  through an explicit dependence upon other signals (e.g.,  $X(t)$ , in which case it may be referred to as a "state-weight").

When  $w(t,s) = w(t-s)$ , Eq. (2) becomes a convolution integral, and the time-weighting function is interpreted as a moving window. If  $T$  is a characteristic time, or duration of  $w$ , the dependence may be expressed as  $w(t,s; T)$ , and a shorthand notation used for the integral

$$\int_{-\infty}^{\infty} w(t-s; T) \dots ds = \langle \dots \rangle_T. \quad (4)$$

Eq. (4) defines a time average on a time scale  $T$ . In the notation,  $w(t,s; T)$ , the independent variable,  $s$ , is used as the domain of the function, the independent variable,  $t$ , represents the current time, and is used to parameterize the choice of  $w$ , enabling the time window to change or move as time changes or evolves. This dependence on  $t$  also allows the user to change the shape of time-weighting as  $t$  changes. Moreover, the time scale can vary with time (i.e.,  $T = T(t)$ ). The weight function may also depend upon other information besides time (e.g., the raw signal, the feature signal, signals derived from the feature signal (such as its time derivatives), other "auxiliary" signals,  $S(t)$ , or control signals,  $U(t)$ ). This generalization allows for state-weighting as mentioned above, or to, e.g., include other "outside" information into the analysis, emphasizing feature signal information accordingly.

$f(t,D)$  of Eq. (2) reduces to the instantaneous density defined by Eq. (1) when  $w(t,s) = \delta(t-s)$ . The notation of Eq. (4) may be used for both continuous and discrete time averages.

For example, the time window may be a rectangular moving window of length  $T$ , such that

$$w(t, s; T) = w(t - s) = \frac{1}{T} \theta[(T - (t - s))(t - s)], \quad (5)$$

where  $\theta(x)$  is the Heaviside step function, it results that

$$f(t, D; w) = \langle \delta[D - X(s)] \rangle_T = \frac{1}{T} \int_{t-T}^t \delta[D - X(s)] ds. \quad (6)$$

In another example, the time window may be an exponential moving window  
5 with time constant  $T$ , such that

$$w(t, s; T) = w(t - s) = \frac{1}{T} \exp\left(\frac{s-t}{T}\right) \theta(t - s), \quad (7)$$

wherein it results that

$$f(t, D; w) = \langle \delta[D - X(s)] \rangle_T = \frac{1}{T} \int_{-\infty}^t \exp\left(\frac{s-t}{T}\right) \delta[D - X(s)] ds. \quad (8)$$

This particular choice of weight is the preferred embodiment for all analog  
10 implementations because of its ease of use.

A time-weighted cumulative distribution function (TWCDF) of the feature  
signal,  $F(t, x; w)$ , can now be derived. The TWCDF describes the distribution of the  
feature signal in a time window that is weighted according to the time-weight  
function,  $w$  (hence the notational dependence of  $F$  on  $w$ ). This time-weight  
15 determines how information from the feature signal will contribute to the probability  
distribution, and is described in more detail below.

It will be appreciated that the TWFD can be obtained from the TWCDF by  
differentiation, and the TWCDF obtained from the TWFD via integration as follows:

$$f(t, x; w) = F_x(t, x; w), \quad (9)$$

$$F(t, x; w) = \int_{-\infty}^x f(t, y; w) dy. \quad (10)$$

It is easy to see from Eq. (10) and the well known identities relating the Dirac  
 $\delta$ -function and Heaviside functions:

$$\theta(x) \equiv \int_{-\infty}^x \delta(s) ds,$$

25 (11)  
and

$$\delta(x) = \frac{d}{dx} \theta(x),$$

(12)

that all equations for feature densities (Eqs. (2), (6), and (8) hold for cumulative distribution as well, provided that symbols " $f$ " and " $\delta$ " in these equations are  
5 replaced by " $F$ " and " $\theta$ ", respectively.

As will be appreciated by one with ordinary skill in the art, percentile values are the very building blocks of probability distributions and enable a robust statistical description thereof. In many applications they produce significantly better information than other more commonly utilized statistics such as the mean and  
10 standard deviation. The set of all percentile values completely describes the distribution from which they are derived.

Given any number  $p \in [0, 1]$ , the  $p^{\text{th}}$  percentile,  $X_p(t)$ , of the TWCDF,  $F(t, x; w)$  is defined implicitly by the equation

$$F(t, X_p(t; w); w) = p.$$

15

(13)

Differentiating the above equation with respect to  $t$ , we see that

$$F_t(t, X_p) + F_x(t, X_p) \dot{X}_p(t) = 0$$

(14)

In the above equation,  $F_t$  and  $F_x$  denote the partial derivatives of  $F$  with  
20 respect to  $t$  and  $X$ , respectively. Rearranging terms results in

$$\dot{X}_p(t) = -\frac{F_t(t, X_p)}{F_x(t, X_p)}.$$

(15)

The denominator of the above equation is the probability density function,  $f(t, X; w)$ , referred to as the "time-weighted feature density." Thus,

$$\dot{X}_p(t) = -\frac{F_t(t, X_p)}{f(t, X_p)}; X_p(0) = X(0).$$

25

(16)

The solution to this important differential equation is the  $p^{\text{th}}$  percentile,  $X_p(t)$ . In other words,  $X_p(t)$  is the  $p^{\text{th}}$  percentile of the time-varying cumulative distribution function,  $F(t, X; w)$ , generated by the variations of the feature signal,  $X(t)$ , in the  
30 temporal window defined by  $w$ .

In most practical applications, the TWCDF will not be known (or immediately available for on-line use) at each moment in time and will instead need to be approximated from available information.

One method of estimating the TWCDF (or the corresponding TWFD) is accomplished by partitioning the state space (containing the range of the feature signal) and computing time-weighted histograms that keep track of how often each bin is visited by the feature signal. This non-parametric approach to obtaining  $\hat{F}(t, x; w)$  has the advantage (over the parametric approach described later herein) of allowing greater flexibility in approximating the TWCDF with any level of precision desired (although as precision improves, complexity of the implementation increases).

The feature density for discrete (digitally sampled) data can be computed in finite differences as follows (suppressing dependence on  $w$ ):

$$\begin{aligned} \hat{f}(t, D_j) &= \frac{1}{D_{j+1} - D_{j-1}} [\hat{F}(t, D_{j+1}) - \hat{F}(t, D_{j-1})] \\ &= \frac{1}{D_{j+1} - D_{j-1}} \langle \theta [D_{j+1} - X(s)] [D_{j-1} - X(s)] \rangle_t \end{aligned} \quad (17)$$

or

$$\hat{f}_{ij} = \hat{f}(t_i, D_j) = \sum_k w_{i-k} \theta_{ij}, \quad (18)$$

where  $w_{i-k} = \frac{1}{2}(t_{i+1} - t_{i-1})w(t_i - t_k)$ , and  $\theta_{ij} = \theta((D_{j+1} - X(t_i))(D_{j-1} - X(t_i)))$ .

When the discrete character of the data is not intrinsic but a result of sampling and digitization of a continuous signal, then some kind of interpolation between the consecutive data points can often be appropriate. Then the identity

$$\delta[a - f(x)] = \sum_i \frac{\delta(x - x_i)}{|f'(x_i)|} \quad (19)$$

can be utilized to compute the feature density. In Eq. (19),  $|f'(x)|$  denotes the absolute value of the function derivative with respect to  $x$  and the sum goes over all

$x_i$  such that  $f(x_i) = a$ . For example, if linear interpolation is adequate, one can use  $w_{i-k} = (t_{i+1} - t_i)w(t_i - t_k)$ , and

$$\theta_{ij} = \begin{cases} 1 & \text{if } D_j \leq X(t_i) < D_{j+1}, \quad D_j \leq X(t_{i+1}) < D_{j+1} \\ \frac{D_{j+1} - D_j}{X(t_{i+1}) - X(t_i)} & \text{if } X(t_i) < D_j, \quad X(t_{i+1}) \geq D_{j+1} \\ \frac{X(t_{i+1}) - D_j}{X(t_{i+1}) - X(t_i)} & \text{if } X(t_i) < D_j, \quad D_j \leq X(t_{i+1}) < D_{j+1} \\ \frac{D_{j+1} - X(t_i)}{X(t_{i+1}) - X(t_i)} & \text{if } D_j \leq X(t_i) < D_{j+1}, \quad X(t_{i+1}) \geq D_{j+1} \\ \frac{D_{j+1} - X(t_{i+1})}{X(t_i) - X(t_{i+1})} & \text{if } X(t_i) \geq D_{j+1}, \quad D_j \leq X(t_{i+1}) < D_{j+1} \\ \frac{X(t_i) - D_j}{X(t_i) - X(t_{i+1})} & \text{if } D_j \leq X(t_i) < D_{j+1}, \quad X(t_{i+1}) < D_j \end{cases}$$

5 (20)

In order to compute the approximation to the cumulative distribution, the following formula may be used:

$$\hat{F}(t, D) = \sum_{i=0}^{N-1} \left[ \hat{F}_i(t) + \left[ \hat{F}_{i+1}(t) - \hat{F}_i(t) \right] \frac{D - D_i}{D_{i+1} - D_i} \right] \theta[(D - D_i)(D_{i+1} - D)],$$

(21)

- 10 where  $\hat{F}_i(t) = \langle \theta[D_i - X(s)] \rangle_T$ , and  $D_0$  and  $D_N$  are such that  $\hat{F}_0(t) = 0$  and  $\hat{F}_N(t) = 1$ , respectively (i.e., selected so that the signal never goes outside the interval  $[D_0, D_N]$ ). It will be appreciated that the cumulative distribution function approximation obtained using by Eq. (21) is a continuous function of  $D$ . It will also be appreciated that, for a given  $s$ ,  $\theta[D_i - X(s)]$  need not be evaluated at each  $D_i$ . Instead, a binary search can
- 15 be used to find  $i$  such that  $D_i \leq X(s) < D_{i+1}$ .

- Another method of estimating the TWCDF (or the corresponding TWFD) is accomplished by approximating  $F(t, x; w)$  by a model distribution function,  $\tilde{F}(t, x; w, v(t))$ , that may depend upon a vector of parameters,  $v(t)$ . Some typical examples include a Gaussian (normal), or chi-squared distribution with
- 20  $v(t) = [\mu(t) \sigma(t)]$ , (the mean and standard deviation); a uniform distribution with  $v(t) = [a(t) b(t)]$  (the left and right endpoints); a triangular distribution with



$v(t) = [a(t) \ b(t) \ c(t)]$  (the left, vertex, and right endpoints of the density, respectively); or an exponential distribution with  $v(t) = \lambda(t)$ , the inverse of the distribution's mean.

An estimate,  $\hat{v}(t)$ , may then be obtained of the parameter vector,  $v(t)$ , from information available up to time  $t$ . This facilitates a "parametric" estimate of

5  $\hat{F}(t, x; w) = \tilde{F}(t, x; w, \hat{v}(t))$ , by substituting the present parameter estimate into the model distribution.

The true feature density is typically unavailable for online analysis but may be well-approximated, assuming the distribution is well-modeled by a parametric distribution, through estimation of the parameters upon which the distribution  
10 depends. Often the moments of the distribution may be computed and used as input to the model to determine the approximation.

For example, a Gaussian approximation can be achieved using the first two moments (the mean,  $\mu$ , and the variance,  $\sigma$ ) of the feature signal. More precisely, the model density

$$15 \quad \tilde{f}(t, x, T, [\mu \ \sigma]) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left[ -\frac{(x - \mu)^2}{2\sigma} \right]$$

(22)

is used, along with the parameter estimators

$$\hat{v}(t) = [\hat{\mu}(t) \ \hat{\sigma}(t)] = [\langle X \rangle_T \ \langle X^2 \rangle_T - (\langle X \rangle_T)^2],$$

(23)

20 to obtain

$$\hat{f}(t, x, T) = \tilde{f}(t, x, T, \hat{v}(t)).$$

(24)

Similarly other parametric approximations for data may be used, such as uniform or triangular distributions. Though the true distribution of the data is not  
25 exactly known, approximations work well because the algorithm tracks the percentile at each instant in real time. The uniform and triangular distributions have the advantage of stability, faster processing, and, most importantly, ease of hardware implementation over others that involve exponentiation. The presence of the exponential function for Gaussian and Chi-squared distributions (for absolutely  
30 positive data) increases computation in software and complexity of hardware, but may also be used. The uniform distribution model reduces the complexity of

implementing circuitry by an order of magnitude. This reduction is significant in reducing the size and power requirement of the circuit and significantly improves scope for miniaturization and possible embedded applications. Further, the uniform distribution makes the implementation stable, unlike the Gaussian approximation which requires signal limiting (clipping) to prevent output from becoming unbounded due to the exponential function. Similarly, the use of a triangular distribution model for data that is strictly positive simplifies the implementation (assuming the left end and vertex of the density are at  $x=0$ , the distribution is completely specified by the right endpoint, i.e., the maximal  $x$  value). The triangular distribution approximation can also be used to estimate the distribution of the signal from a single percentile calculation. One with ordinary skill in the art will appreciate that similar parametric modeling and estimation may be performed using any other distribution model and parameter estimation scheme meant to approximate the underlying true density/distribution.

15

#### Step 4: Compute the Percentile Tracking Filter (PTF)

Once the TWCDF estimate,  $\hat{F}(t, x; w)$ , has been obtained (and the TWFD estimate,  $\hat{f}(t, x; w)$ , is therefore also available), an estimator of  $X_p(t)$ , referable to as the output of a percentile tracking filter (PTF) at time  $t$ , and denoted by  $\hat{X}_p(t)$ , is obtained, as shown in box 28, from the differential equation

$$\dot{\hat{X}}_p(t) = -\frac{\hat{F}_t(t, \hat{X}_p)}{\hat{f}(t, \hat{X}_p)}; \hat{X}_p(0) = X(0).$$

(25)

One with ordinary skill in the art will appreciate that  $p$  may denote a vector of percentile values, (i.e.,  $p = [p_1 \ p_2 \ \dots \ p_n]$ ) in which case  $\hat{X}_p(t)$  denotes the vector  $[\hat{X}_{p_1}(t) \ \hat{X}_{p_2}(t) \ \dots \ \hat{X}_{p_n}(t)]$ .

25

In certain applications, it is desirable to know the value of  $X_p(t)$  such that  $F(t, X_p; w) = p = \text{constant}$ . Thus,  $X_p(t)$  is an output of a rank-order (also order statistic, quantile, or percentile) filter. For example,  $X_{1/2}(t)$  is the output of a median

filter, producing at each moment in time the median of the  $w$ -weighted distribution of feature signal information in the most recent window.

Numerical rank-order filtering is a computationally expensive operation. First, it requires knowing, at any given time, the values of  $N$  latest data points, where  $N$  is the length of the moving window, and the numerical and chronological order of these data points. This memory requirement is a major obstacle to implementing an analog order statistic filter. Another computational burden on different (numerical) rank-order filtering algorithms results from the necessity to update the numerically ordered/sorted list, i.e., to conduct a search. Overall performance of a rank-order filter is a trade-off between performance, speed, and memory requirements.

$\hat{X}_p(t)$  is the  $p^{\text{th}}$  percentile of  $\hat{F}(t, x; w)$  in a moving window  $w(t, s; T) = w(t - s; T)$  if

$$\hat{F}(t, \hat{X}_p; T) = \int_{-\infty}^{\hat{X}_p} \hat{f}(t, D; w) dD = \int_{-\infty}^{\infty} w(t - s; T) \theta[\hat{X}_p(t) - X(s)] ds = p.$$

(26)

where  $0 \leq p \leq 1$  is the percentile value. The time derivative of  $\hat{X}_p$  can now be computed as

$$\dot{\hat{X}}_p(t) = -\frac{\hat{F}_t(t, \hat{X}_p; w)}{\hat{f}(t, \hat{X}_p; w)} = -\frac{1}{\hat{f}(t, \hat{X}_p; w)} \int_{-\infty}^{\infty} \dot{w}(t - s; T) \theta[\hat{X}_p - X(s)] ds.$$

(27)

Typically  $w(\cdot)$  is chosen such that it vanishes at  $\pm \infty$ , and since  $\dot{w}(t - s) = -\frac{d}{ds} w(t - s)$ ,

integration of Eq. (27) by parts leads to

$$\dot{\hat{X}}_p(t) = \frac{1}{\hat{f}(t, \hat{X}_p; w)} \int_{-\infty}^{\infty} w(t - s; T) \left\{ \dot{X}(s) \theta[\hat{X}_p(s) - X(s)] \right\} ds.$$

(28)

$$= \frac{1}{\hat{f}(t, \hat{X}_p; w)} \sum_i w(t - t_i) \text{sign} \left[ \dot{X}(t_i) \right],$$

(29)

where the sum goes over all  $t_i$  such that  $X(t_i) = X_p(t)$ , i.e., over all times of threshold crossings, using the identity of Eq. (19).

It will be appreciated that if  $w(\cdot)$  has a shape that coincides with the impulse response of an analog filter, then the differential equation Eq. (27) can be solved in an analog circuit, provided that  $\hat{f}(t, x; w)$  is evaluated in finite differences according to Eq. (17).

For example, where the time-weight function is a rectangular moving window of length  $T$ ,

$$\dot{\hat{X}}_p = \frac{1}{T \hat{f}(t, \hat{X}_p; T)} \left\{ \theta[\hat{X}_p - X(t-T)] - \theta[\hat{X}_p - X(t)] \right\}. \quad (30)$$

The braced expression is merely the total number of upward crossings of the threshold,  $\hat{X}_p(t)$ , by the signal minus the total number of downward crossings, in agreement with Eq. (29).

In another example, where the time-weight function is an exponential moving window of length  $T$ ,

$$\dot{\hat{X}}_p = \frac{1}{T \hat{f}(t, \hat{X}_p; T)} \left\{ p - \theta[\hat{X}_p - X(t)] \right\}. \quad (31)$$

In this example  $p$  enters the equation for  $\hat{X}_p(t)$  explicitly rather than through the initial condition only, as in a general case of Eq. (27). The exponential window is causal, and has an advantageously low computational cost (including low memory requirements) and easy analog implementation. For these reasons this choice of weight function is preferred.

From the PTF output signal,  $\hat{X}_p(t)$ , can be obtained another approximation, denoted as  $\hat{F}(t, x; w)$ , for the TWCDF (and TWFD) by standard interpolation or extrapolation means (e.g., linear interpolation and extrapolation to enable evaluation of the distribution function approximation at values between ordered pairs  $\{(\hat{X}_{p_i}(t), p), i = 1 \dots n\}$  that serve as nodes for the distribution approximation).

The PTF output (and the TWCDF estimates) make accessible important information about the level of quantified features present in the specialized time-

window defined by the time-weight function. Prior to this invention, none of this information could be obtained in a completely analog system, and the computational cost of deriving this information in digital implementations was significantly more expensive and less efficient.

5

Step 5: (Optional) Normalize the Signal

The present invention facilitates an optional normalization technique which may be performed at this point, as shown in box 30. The desirability of performing this optional step will depend upon the nature of the signals of interest, particularly  
10 their respective scales.

Given any input signal,  $z(t)$ , and any (possibly time-varying) cumulative distribution function,  $F(t,x)$ , a new signal,  $y(t)$ , may be obtained by evaluating the distribution function at the input signal value, i.e.,  $y(t) = F(t, z(t))$ . This new signal may be referred to as the normalization of signal  $z$  with respect to the distribution  $F$ .

15 The signal,  $y(t)$ , is deemed "normalized" because, no matter what the values/range taken by the input signal, the resulting signal values are always in the interval  $[0,1]$ . The flexibility of this normalization technique is that any distribution function and input signal may be utilized (provided only that  $F$  is defined so that its second argument is of the same dimension as the input signal), as well as the  
20 observation that, for a stationary distribution (i.e.,  $F(t,x) = F(x)$ ) the procedure is a monotonic transformation of the input signal, make this a useful tool for signal analysis.

This normalization technique can be combined with the herein described techniques for approximating the TWCDF of a signal, to enable normalization of an  
25 input signal with respect to a time-weighted distribution function of the feature signal of the same (or a different) signal. More specifically, the combination of methods results in the normalized signals  $\hat{y}(t) = \hat{F}(t, x(t); w)$  and  $\hat{\hat{y}}(t) = \hat{\hat{F}}(t, x(t); w)$ .

This normalization technique (together with the method for detecting changes in a time-varying distribution function described elsewhere herein) has been  
30 successfully applied, for example, to the problem of automated speech recognition, accurately detecting each occurrence of a particular phoneme in digitally recorded speech.

Step 6: Compare Foreground and Background or Reference Distributions/Densities  
in Order to Detect and Quantify Changes in Signal Features

Referring to FIG. 2, the foreground signal and background or reference signal may now be compared in order to detect or quantify feature changes in the foreground signal. That is, the ability to extract information from (or restrict the influence of information to) different time scales through weighting functions, and the ability to precisely control the set of features under study, allows further use of feature density analysis method as a component of a system for detection and quantification of feature changes. Referring also to FIG. 3, such detection and quantification is accomplished by comparing the PTF outputs (or entire time-weighted feature densities) in the moving foreground window with that of a background or established reference from which a specified change is to be detected.

The method described thus far makes available TWCDF approximates and associated PTF signals for various percentiles. These are determined in part by the time-weight function used in their definition and computation, which describes the way information is weighted and utilized in the production of these approximations. These approximations can be further analyzed and utilized to produce new and highly valuable means for detection and quantification of changes in the feature signal, and therefore in the underlying system that provided the raw signal input.

Generally, the concept is one of comparison between results obtained for different time-weighting of the information provided by the percentile tracking filter outputs or approximations of the TWFD or corresponding TWCDF of the feature signal. One skilled in the art will appreciate that the user may specify two (or more) time-weight functions,  $w_1$  and  $w_2$ , with different characteristic timescales,  $T_1$  and  $T_2$ , respectively, and thereby obtain two sets of PTF outputs and TWCDF approximations:

$$\hat{X}_p(t; w_1) \text{ and } \hat{F}_1(t, x; w_1), \hat{\hat{F}}_1(t, x; w_1) \text{ corresponding to timescale } T_1, \text{ and}$$

$$\hat{X}_p(t; w_2) \text{ and } \hat{F}_2(t, x; w_2), \hat{\hat{F}}_2(t, x; w_2) \text{ corresponding to timescale } T_2.$$

$T_1$  is preferably chosen to be much larger than  $T_2$  so that a comparison may be performed of the above quantities, interpreting the former as representing the

background or reference information (i.e., obtained from a large time window of past or historical information) and the latter set as being representative of the foreground or more current, test information (i.e., obtained from a small time window of recent information). The existence of a reference for the comparison allows a built-in type

5 of normalization that ensures comparison of "apples to apples" in the resulting analysis. The reference information need not necessarily be continually updated (the time between updates could, e.g., be proportional to the timescale analyzed), or a constant set of information (e.g., a constant,  $C$ , and fixed distribution,  $F_{ref}(t,x)$ ) may be used as a reference for the comparison with

10  $\hat{X}_p(t; w_2)$  and  $\hat{F}_2(t, x; w_2)$  or  $\hat{F}_2(t, x; w_2)$  respectively. One skilled in the art will recognize that any of the standard techniques for statistical analysis of data and for comparison of distributions or densities can also now be applied for time-dependent quantification of the signal and its changes.

The wealth of comparisons that may be performed between the PTF outputs

15 are well known to those skilled in the art of signal analysis. Although such outputs were not available prior to the present invention, there are numerous known techniques for comparison of test signals to reference signals that may be employed to these newly available PTF outputs. A typical example (utilizing the type of comparison employed in the successful detection algorithm disclosed in U.S. Patent

20 No. 5,995,868) involves computing a ratio of the foreground and background outputs, e.g.,

$$r(t) = \frac{\hat{X}_p(t; w_2)}{\hat{X}_p(t; w_1)},$$

(33)

and then comparing the resulting ratio to a threshold or reference value that, when

25 reached, signifies occurrence of the change that was to be detected.

Comparisons between test and reference time-weighted densities and distributions of the feature signal have been invented to allow the user to quantify differences and detect a wide array of changes in these distributions with great sensitivity and specificity. Further, these comparisons (and those mentioned above

30 for the PTF) have also been implemented in analog form and in highly efficient digital form.

A time-varying signal, referred to as a  $\Lambda$ -estimator, is defined to quantify the difference between distributions by functioning as a type of distance measure between them as

$$\begin{aligned}\Lambda(t; W, G) &= \Lambda(F_1(t, x; w), F_2(t, x; w)) \\ &= \int_{-\infty}^{\infty} W(F_1(t, y; w)) G(F_1(t, y; w) - F_2(t, y; w)) dF_1(y) \\ &= \int_{-\infty}^{\infty} W(F_1(t, y; w)) G(F_1(t, y; w) - F_2(t, y; w)) f_1(y) dy\end{aligned}$$

5 (34)

for some percentile weighting function,  $W$ , and some spatial weighting function,  $G$ .

The  $\Lambda$ -estimator, through choice of weighting functions, is able to quantify a wide array of differences in the two distributions being compared.

For example, consider an estimator of the differences between the foreground  
10 and background distributions as follows:

$$\Lambda = \lambda_{\pm} = \int_{-\infty}^{\infty} W[F_{bg}(x)] [F_{bg}(x) - F_{fg}(x)] \theta\{\pm [F_{bg}(x) - F_{fg}(x)]\} dF_{bg}(x),$$

(35)

where the percentile weighting function  $W(z)$  is such that  $\int_0^1 z W(z) dz = 1$ , and a pair of  
spatial weighting functions (giving rise to a pair of estimators,  $\lambda_{\pm}(t)$ , resp.) is given  
15 by  $G_{\pm}(x) = x \theta(\pm x)$ . The above parameters compute weighted distance between the  
background and the foreground for specific cases of  $F_{fg} < F_{bg}$  and  $F_{fg} > F_{bg}$  for  $\lambda_+$  and  
 $\lambda_-$  respectively.

Computing  $\lambda_{\pm}$  estimators becomes especially easy if the weighting function  
 $W(x)$  is a delta function,  $W(x) = \frac{1}{p} \delta(x - p)$ . Thus Eq. (35) reduces to

$$\lambda_{\pm} = \left[ 1 - \frac{1}{p} F_{fg}(x_p) \right] \theta \left\{ \pm \left[ 1 - \frac{1}{p} F_{fg}(x_p) \right] \right\},$$

(36)

where  $x_p$  is such that  $F_{bg}(x_p) = p$ , i.e., the output of the PTF for  $p^{\text{th}}$  percentile. Using  
the fact that  $F_{fg}[x_p(t)] = \langle \theta[x_p(t) - X(s)] \rangle_{T_R}$ , Eq. (36) becomes



$$\lambda_{\pm}(t) = \frac{1}{p} \left[ p - \langle \theta[x_p(s) - X(s)] \rangle_{T_{\#}} \right] \theta \left\{ \pm \left[ p - \langle \theta[x_p(s) - X(s)] \rangle_{T_{\#}} \right] \right\} \quad (37)$$

where we replaced  $x_p(t)$  by  $x_p(s)$  in the assumption that  $T_{bg} \gg T_{fg}$ , and therefore  $x_p(t)$  changes slowly with respect to  $X(t)$ .

5 For example, consider a simple mix of two signals with  $W(x) = 2\delta\left(x - \frac{1}{2}\right)$ ,

$F_{bg}(x) = F(x)$ , and  $F_{fg}(x) = (1 - \alpha)F(x) + \alpha F\left(\frac{x}{q}\right)$ . That is, the foreground window

contains a mix of two signals such that the ratio of their mean values is  $q$ . Then

$$\lambda_{\pm} = \alpha \left[ 1 - 2F\left(\frac{x_{y_2}}{q}\right) \right] \theta[\pm(q-1)],$$

(38)

10 where  $x_{y_2}$  is such that  $F\left(x_{y_2}\right) = \frac{1}{2}$ , i.e., the background median. If  $\lim_{x \rightarrow 0} F(x) = 0$ ,

and  $q + \frac{1}{q} \gg 1$ , then  $1 - 2F\left(\frac{x_{y_2}}{q}\right) \approx \text{sign}(q-1)$  and

$$\lambda_{\pm} \approx \pm \alpha \theta[\pm(q-1)].$$

(39)

15 Thus the values of  $\lambda_{\pm}$  become a measure of fraction of the second component in the foreground window.

The invariance of the  $\Lambda$  estimators given by Eq. (34) to a monotonic transformation (function) of the signal  $f[X(t)]$ , e.g.,  $a > b \Rightarrow f(a) > f(b)$ , allows simplification of computations by rescaling the signal to a convenient range, and  
 20 makes the change detection invariant to non-linear amplification of the raw signal. In combination with a signal preprocessing system which provides methods for taking a raw input signal, measuring desired features or properties of that signal, then transforming the output into (several) signals of amplitudes which characterize the levels of the particular features under study as they evolve over time, the above  
 25 mentioned tools enable rapid, efficient, and reliable detection of changes in these specific features and tracking of their time evolution.

In both FIGs. 2 and 3, thresholding of the resulting comparison output is performed, as shown in box 32, in order to detect feature changes exceeding a pre-defined value. Where such feature changes are detected, an alarm, whether visual, audible, or otherwise, may be communicated, as shown in box 34.

5

#### Step 7: (Optional) Predict Future Changes

Lastly, the present invention's ability to rapidly and accurately detect changes in certain features of an input signal can be used to predict future changes in cases when the detected changes are associated with an increased likelihood of these future changes. For example, when applied to seismic signals, the method can enable prediction of an earthquake or volcanic eruption; when applied to meteorological signals, the method can enable prediction of severe weather; when applied to financial data, the method can enable prediction of an impending price change in a stock; when applied to brain waves or heart signals, the method can enable prediction of an epileptic seizure or ventricular fibrillation; and when applied to brain wave or electromyographic signals, the method can enable prediction of movement of a body part.

#### Example: Financial Analysis

In an exemplary illustration of application of the present invention to financial analysis, a three-dimensional raw input signal,  $x(t)$ , is received such that, as shown in FIG. 4,  $x(t)$  is the price at time  $t$  of three stocks (IBM, GM, and MRK) during a 360 minute period on October 23, 1994.

Pre-processing is then performed on  $x(t)$  to yield a feature signal,  $X(t)$ , as shown in FIG. 5, such that  $X(t)$  is the value of a portfolio at time  $t$  comprising the three stocks of interest:  $X(t) = 300*GM + 100*IBM + 80*MRK$ .

Next, as time,  $t$ , evolves, the time-weight function,  $w(t,s)$ , is defined for use at each time point. In the present example,  $w$  is defined so that the time-weight is a moving square window of length  $T=100$  minutes. In practice, more complicated time-weights could be used that include a dependence upon an auxiliary signal such as oil prices and supplies which may have a direct impact in GM stock prices by influencing consumer preferences and cash availability. Another time-weight, for example is the average temperature across the U.S. for the past week (thus incorporating information that could have an effect on the feature signal if, e.g., a

freeze altered the price of wheat futures or heat wave altered the price of California energy companies) that could directly or indirectly affect these or other stocks. The true 0.25, 0.50, and 0.75 percentiles are shown in FIG. 6 as, respectively,  $X_{0.25}(t)$ ,  $X_{0.50}(t)$ , and  $X_{0.75}(t)$ . The true feature density,  $f(t, w)$ , and distribution,  $F(t, w)$ , are  
 5 calculated from the data at two different times,  $t_1$  and  $t_2$ , as shown in FIGs. 7 and 8. The parameters are calculated for a time-scale of 100 minutes.

Next, the feature signal and time-weight are used to compute/update an evolving approximation to the time-weighted density and corresponding distribution function of the feature signal:  $\hat{f}(t, x; w)$  and  $\hat{F}(t, x; w)$ . In the present example,  
 10  $\hat{f}(t, x; w_1)$  and  $\hat{F}(t, x; w_1)$  are evaluated assuming a Gaussian (normal) density approximation (bell-shaped) for the data over the past 100 minutes. These approximations can be compared to the true distributions shown in the previous figure. Again, these were evaluated at two different times,  $t_1$  and  $t_2$ , as shown in FIGs. 9 and 10.

15 The feature signal, a specified set of one or more percentile values,  $p$ , and approximations  $\hat{f}(t, x; w)$  and  $\hat{F}(t, x; w)$  are then used to compute/update the PTF output,  $\hat{X}_p(t; w)$ . It will be appreciated that  $p$  may be a vector of multiple percentiles, as in this example where  $p = [0.25, 0.50, 0.75]$ . The PTF output,  $\hat{X}_p(t; w)$ , is shown in FIG. 11.

20 An interpolation/extrapolation scheme is then used to compute/update a second set of (evolving) approximations to the time-weighted density and corresponding distribution function of the feature signal,  $\hat{\hat{f}}(t, x; w)$  and  $\hat{\hat{F}}(t, x; w)$ . In the present example,  $\hat{\hat{f}}(t, x; w_1)$  and  $\hat{\hat{F}}(t, x; w_1)$  were determined using the outputs of the PTF by linear interpolation. These were again evaluated at the times,  $t_1$  and  $t_2$ , as  
 25 shown in FIGs. 12 and 13.

Then the PTF output and approximations to the time-weighted density/distribution of the feature signal are analyzed to detect, quantify, or predict changes in the system that produced the raw signal. This analysis may consist, for example, of establishing or computing a reference against which to compare the  
 30 information being generated. One preferred approach is to use a fixed reference value and a fixed density/distribution and compare them to the PTF and the

density/distribution approximations, respectively. A second preferred approach involves performing the prior method steps simultaneously with two differing choices of time-weight function, one to establish a reference PTF and density/distribution approximation and the other to generate a test PTF and density/distribution approximation, then comparing the two resulting sets of information. Described above is a method for comparing two PTF outputs (e.g., computing their ratio) and for comparing test-to-reference distributions ( $\Lambda$ -estimators). The ratio and/or  $\Lambda$ -estimators are used to compare the feature content in one time-window/scale to another. Changes may be detected, e.g., by applying thresholds to either ratio and/or  $\Lambda$ -estimators. In the present example, the above calculations were performed on short time-scale of 100 min. The long time-scale calculations were performed using the portfolio variations starting from 1980 until 2001 with a time-scale of  $T=1$  year. This was taken as reference for the calculation of the  $\Lambda$  parameter. Thus, a simple difference between  $\hat{F}(t)$ , evaluated for a  $T = 100$  min and  $\hat{F}_{REF}$ , evaluated with a time-scale of  $T=1$  year was used for  $\Lambda(t)$ , as shown in FIG. 14.

It should be noted that the  $\Lambda(t)$  signal on the 100 minute time scale was unaffected by the more brief outlier spike that occurred in the feature signal approximately 260 minutes into the day's trading.

Lastly, once features of the financial system from which the raw signal was measured have been detected and quantified, the output may be utilized. One such use of this information is signal normalization (of any signal) with respect to any of the distribution approximations made available by the present invention. Another use involves prediction of future changes if it happens that the detected changes are associated with an increased likelihood of certain future signal changes.

In the present example, the portfolio value on the particular day did not demonstrate any significant change over the background distribution, so the owner may have simply decided to maintain his holdings at that time.

### System

FIGs. 15 and 17 show block diagrams of a preferred analog system 100 operable to implement the above-described method. FIGs. 16 and 18 show detailed circuit schematics of a preferred embodiment of the analog system 100 shown described generally in FIGs 15 and 17. The analog system 100 comprises

two major components, an analog PTF circuit 102 and an analog Lambda circuit 104. This implementation is based upon a recognition that the use of approximations in determining feature density facilitate tracking of the percentile in real-time with minimal errors. The use of a uniform density distribution makes the implementation simpler by an order of magnitude over implementations involving distributions with exponentiation, while providing an output that closely tracks the actual median obtained by sorting. The implementation outputs the percentile value of the input signal in real time while retaining the controllability and flexibility of a digital algorithm.

Referring first to FIG. 15, the analog PTF circuit 102 broadly includes a peak detector stage 108; a comparator stage 110; a scaling and shifting amplifier stage 112; an adder stage 114; a multiplier and divider stage 116; and an integrator stage 118.

The peak detector stage 108 determines the peak of an incoming signal. For a uniform distribution, the density function  $f(t, w)$  is given by  $1/(b-a)$ , where  $b(t)$  and  $a(t)$  are the maximum and the minimum of the signal respectively. Thus, for  $a = 0$  (this can be achieved using the optional preprocessing step), the peak detector stage 108 gives  $1/\hat{f}$ .

The comparator stage 110 performs the step of computing:  $\theta[X(s) - X_p(s)]$ . The output of the comparator stage 110 is a voltage equal to a saturation voltage ( $\pm 15$  V) of an amplifier component of the comparator stage 110.

The scaling and shifting amplifier stage 112 brings the voltage to 0-10 V. This enables direct subtraction from the input  $10^*p$  in the adder 114.

The adder stage 114 subtracts the output of the scaling and shifting amplifier stage 112 from  $10^*p$ . Thus the output of the adder stage 114 is  $10^*(p - \theta[X(s) - X_p(s)])$ .

The multiplier and divider stage 116 multiplies the output of the adder stage 114 and divides the result by 10. Thus the output of this stage 116 is  $\frac{1}{f}(p - \theta[X(s) - X_p(s)])$ , which is  $TX_p(t)$ .

The integrator stage 118 has an input-output relationship defined by the following equation:  $V_o = \frac{1}{RC} \int_{-\infty}^t V_i dt$ , where  $V_i$  and  $V_o$  are the input and the output

respectively. Thus the output of this stage 118 is  $\frac{1}{RC}TX_p(t)$ . Choosing proper values for  $R$  and  $C$  will give us an output that is equal to  $X_p(t)$ .

Referring also to FIG. 16, the peak detector stage 108 has a natural exponential forgetting factor due to discharging of the capacitor  $C_p$ . The time-  
 5 window of exponential forgetting for the peak can be controlled by varying the resistor  $R_p$ . The time-window over which the percentile is calculated, is controlled by the integrator stage 118. Resistor  $R_i$  and a capacitor  $C_i$  of the integrator stage 110 control the time factor  $T$ . By tuning these parameters,  $R_p$ ,  $C_p$ ,  $R_i$  and  $C_i$ , the PTF circuit 102 can be tuned to approximate a percentile filter with specific properties.  
 10 For example, by choosing the following values:  $R_p = 1.0\text{M}\Omega$ ,  $C_p = 2.2\mu\text{F}$ ,  $R_i = 1.0\text{M}\Omega$ , and  $C_i = 2.2\mu\text{F}$ , there is achieved an exponential forgetting time constant of 2.2 seconds and a percentile time-window of 2.2 seconds.

The output of the PTF circuit 102 with these parameters to an example feature signal input (shown in FIG. 17A) is shown in FIG. 17B. Also shown in FIG.  
 15 17B is the true median obtained by performing a heap sort of sliding 2.2-second window on the input data. Note that the PTF circuit 102 responds to changes in the signal faster than the true median. This is because the true  $N$ -sample median does not respond until  $N/2$  (or  $[N+1]/2$  if  $N$  is odd) samples have passed through the filter. However, due to the implementation of the PTF using exponential forgetting, it  
 20 responds to changes almost instantaneously. This property of the PTF is valuable when attempting to detect rapid signal amplitude changes.

Referring to FIG. 18, the analog Lambda circuit 104 broadly comprises a bank 122 of PTF circuits 102; a bank 124 of reference signals; an adder stage 126; a thresholding stage 128; and an alarm stage 130.

25 The bank 122 of PTF circuits 102 comprises one or more PTF circuits 102, each applied on the original signal,  $x(t)$  or feature signal  $X(t)$ . The output of this bank 122 is the calculation of  $F_{fg}$  (Eq. 35).

The bank 124 of reference signals can be PTF circuits 102 applied to the original raw input or to the feature signal. In such a case, the reference signals are  
 30 typically generated by using large time-scales during the integration step of the PTF operation. Otherwise, the reference signals can be simple constant voltage sources. There is a reference signal corresponding to each of the PTF circuits 102 in the above bank 122. The output of this bank 124 is the  $F_{bg}$ .

The adder stage 126 adds the output of the bank 122 of PTF circuits 102 to the negative output of the bank 124 of reference signals. This stage 126 can also selectively amplify or attenuate the result of each subtraction. Thus the output of the adder stage 126 is a summation of the weighted difference between the PTF circuit  
 5 102 output signal and the reference signal, which is the Lambda parameter:

$$\Lambda(t) = \sum w(t, x)(F_{fg}(t, x) - F_{bg}(t, x)).$$

(40)

$w(t, x)$  can be selected to give a parameter of choice.

The thresholding stage 128 detects increases in the Lambda parameter  
 10 beyond a particular threshold. This stage 128 may be implemented as a comparator.

The alarm stage 130 communicates threshold crossing detected by the thresholding stage 128, and can be implemented as in many different ways, from a simple light (or beep) to a complicated circuit that can perform a sequence of steps, such as a computer that can perform some pre-determined operation.

Referring also to FIG. 19, an analog Lambda circuit is 104 shown operable to  
 15 compute the  $\Lambda$  parameters from the percentiles of the input signal and constant reference signals. The inputs to the analog Lambda circuit 104 are the outputs PTF1, PTF2, and PTF3 provided by parallel PTF circuits 102. PTF1, PTF2, and PTF3 correspond to the  $p_1$ ,  $p_2$  and  $p_3$  percentile of the input signal such that  $p_1 > p_2 >$   
 20  $p_3$ . The reference signals in the circuit 104 are  $Vr_1$ ,  $Vr_2$ , and  $Vr_3$  such that  $Vr_1 > Vr_2 > Vr_3$ . The Lambda circuit 104 comprises first and second stages 134, 136, wherein the first stage 134 is operable to provide the calculated  $\Lambda$  parameter, and the second stage 136 is operable to perform threshold detection and alarm.

The following equations hold for the Lambda circuit 104 shown:

$$V_A = R_p \left( 1 + \frac{R_6}{R_5} \right) \left[ \frac{PTF_1 - Vr_1}{R_9} + \frac{PTF_2 - Vr_2}{R_8} + \frac{PTF_3 - Vr_3}{R_7} \right],$$

(41)

Where

$$R_p = \frac{R_7 R_8 R_9}{R_7 R_8 + R_8 R_9 + R_7 R_9}.$$

(42)

The Lambda circuit 104 can be tuned to output a particular parameter by  
 30 changing the values of resistances  $R_5$ - $R_9$ . For example, if the circuit 104 is to

respond to changes in the  $p_1$  percentile (say the 75<sup>th</sup> percentile), then the gain of that particular input line can be increased while the other two input lines can be attenuated. In the second stage 136, this voltage can be compared to a specific threshold and an alarm can be triggered.

5        The present invention, as described herein, provides computationally efficient characterization of temporally-evolving densities and distributions of signal features of arbitrary-type signals in a moving time window by tracking output of order statistic (e.g., percentile, quantile, rank-order) filters. As noted, the present invention's ability to rapidly and accurately detect changes in certain features of an input signal can  
10 also enable prediction in cases when the detected changes are associated with an increased likelihood of certain of future signal changes.

Although the invention has been described with reference to the preferred embodiment illustrated in the attached drawings, it is noted that equivalents may be employed and substitutions made herein without departing from the scope of the  
15 invention as recited in the claims. Those skilled in the art will appreciate, for example, that although described herein for use with a continuous-time signal, time-weighting can be equally well applied to a discrete-time signal. Thus, the present invention can be implemented in both digital and analog forms.

One skilled in the relevant art will also recognize that any sequence of data,  
20 one- or multi-dimensional, resulting from any measurement, may be used as a signal in the present invention. This extends the applicability of the invention beyond the most commonly used definition of a signal, in which the data is parameterized by an index,  $t$ , that is interpreted as a time variable, to allow  $t$  to be any element of an ordered set. In this event,  $t$  is interpreted simply as an index that determines the  
25 order in which the data appears in the list,  $\{x(t)\}$ , or its structure. This enhances the understanding of how the invention may be utilized in the analysis of lists or sequences. Specific illustrative examples include analysis of word lists such as those found in a document, enabling the quantification of document content, useful, e.g., in electronic search engine applications; as well as analysis of biologic  
30 sequences, structures, or compounds at the molecular, microscopic, or macroscopic level, useful, e.g., in comparison of genetic makeup of a sample to a reference. Similar applications exist to non-biologic sequences, structures or compounds.



Having thus described the preferred embodiment of the invention, what is claimed as new and desired to be protected by Letters Patent includes the following:

5

10

15

20

25

30

## CLAIMS:

1. A method of detecting changes in a signal, the method comprising the steps of:
  - (a) estimating a time-weighted distribution function of the feature signal;
  - 5 (b) defining a percentile tracking filter and producing a PTF output based upon the estimated time-weighted distribution function;
  - (c) comparing the PTF output with a background time-weighted feature distribution function to detect a change in the signal; and
  - (d) communicating detection of the change in the signal.
- 10 2. The method as set forth in claim 1, wherein the signal is of arbitrary type, origin, scale, and degree of complexity.
- 15 3. The method as set forth in claim 1, wherein the signal is selected from the group consisting of the following: continuous-time signals, discrete-time signals, analog signals, digital signals, scalar signals, multi-dimensional signals, deterministic signals, stochastic signals, stationary signals, non-stationary signals, linear signals, nonlinear signals.
- 20 4. The method as set forth in claim 1, wherein the signal is selected from the group consisting of the following: biological signals, financial signals, physical signals, communication signals, mechanical signals, chemical signals.
- 25 5. The method as set forth in claim 1, wherein accomplishing step (b) involves first determining a time-weighted density function of the feature signal.
- 30 6. The method as set forth in claim 1, wherein step (a) is accomplished non-parametrically by partitioning a state space (containing a range of the feature signal) and computing time-weighted histograms that keep track of how often each bin is visited by the feature signal.
7. The method as set forth in claim 1, wherein step (a) is accomplished parametrically.

8. The method as set forth in claim 1, wherein step (a) is accomplished non-parametrically.

9. The method as set forth in claim 1, further including the step of  
5 preprocessing the signal to derive a feature signal prior to performing step (a).

10. The method as set forth in claim 1, further including the step of normalizing the feature signal with respect to the time-weighted distribution function computed in step (a).

10

11. The method as set forth in claim 1, further including step (e) using the detected change in the signal to predict a future change in the signal.

12. The method as set forth in claim 1, further comprising step (e) using  
15 the detected change in the signal to control a process.

20

25

30

13. A system for detecting changes in a signal, the system comprising:  
a PTF circuit operable to function as a percentile tracking filter and generate a  
PTF output signal; and  
a comparator operable to compare the PTF output signal with a reference  
5 signal in order to detect a change in the signal.

14. The system as set forth in claim 13, wherein the PTF circuit comprises:  
a parameter estimation stage operable to compute estimates of parameters  
upon which a model density depends;  
10 a comparator stage operable to determine whether its input exceeds the  
present circuit output;  
an adder stage operable to compute the difference between (the previous  
stage output) and a constant input (representing the percentile to be  
tracked);  
15 a multiplier and divider stage operable to the rate of change of the percentile  
tracking filter output; and  
an integrator stage operable to said rate of change into the circuit output.

15. A system for detecting changes in a signal, the system comprising:  
20 a circuit operable to compute a time-weighted distribution; and  
a Lambda circuit operable to compare the time-weighted distribution with a  
reference signal in order to identify a change in the signal.

16. The system as set forth in claim 15, wherein the Lambda circuit  
25 comprises:  
an adder stage for operable to add the first input to an inverted version of the  
second input so as to generate a Lambda parameter which is a  
summation of a weighted difference between the first and second  
inputs; and  
30 a thresholding stage operable to detect an increase in the Lambda parameter  
beyond a predefined threshold.

17. The system as set forth in claim 16, further comprising an output stage operable to communicate any change in the Lambda parameter detected by the thresholding stage.

5

10

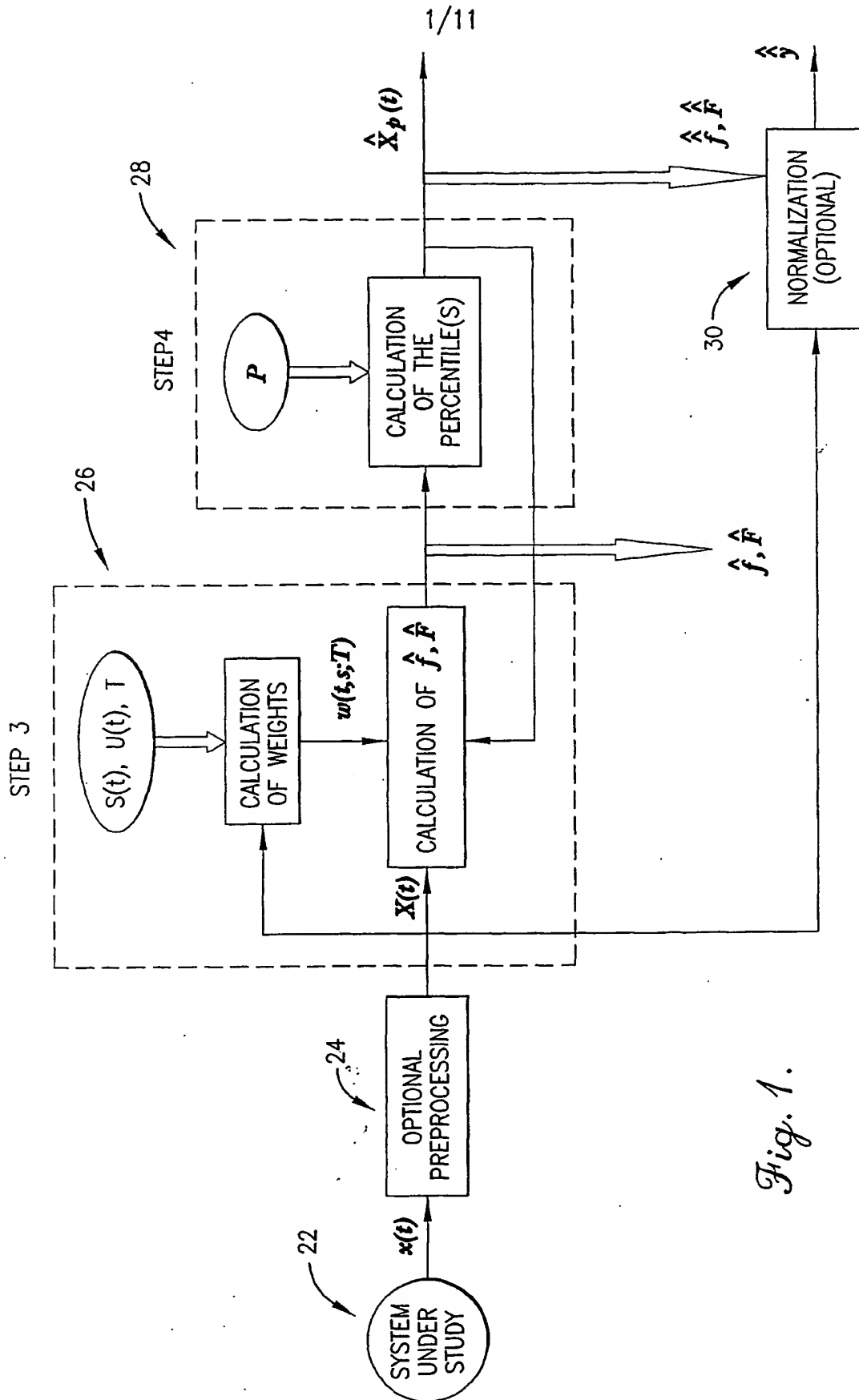
15

20

25

30

35



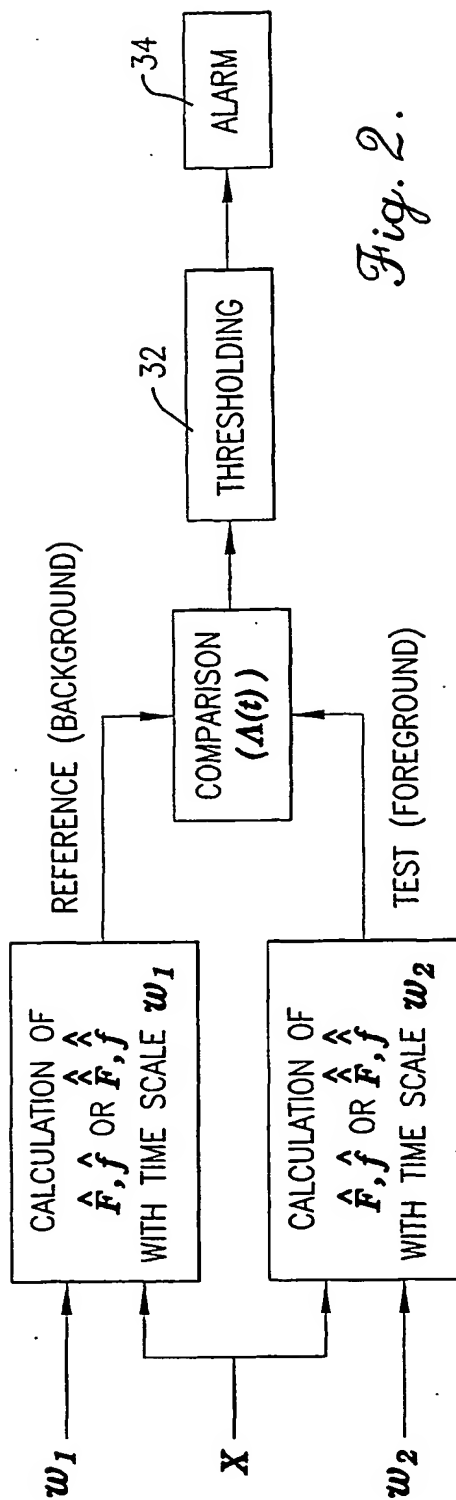


Fig. 2.

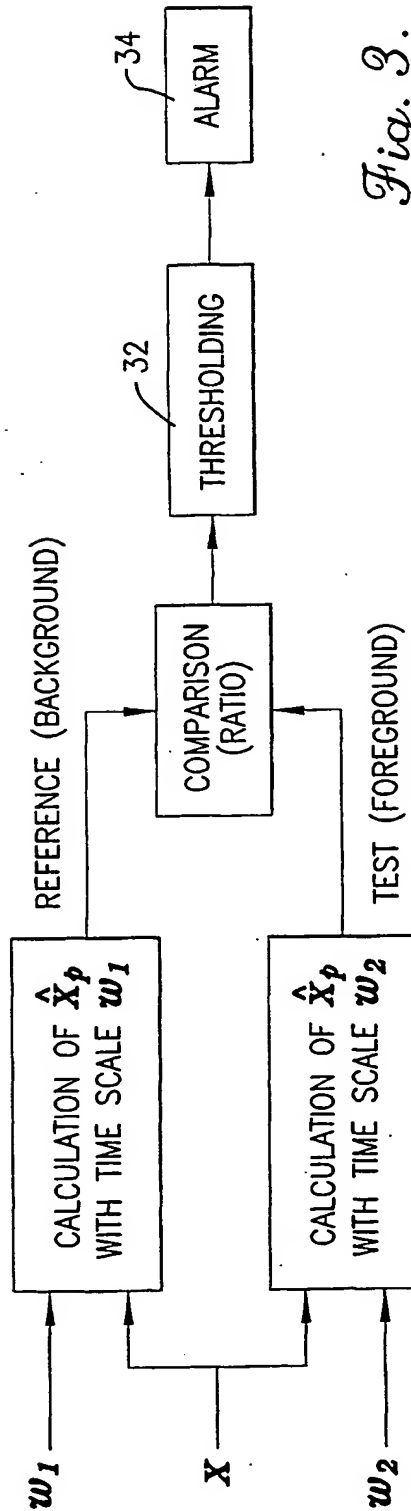
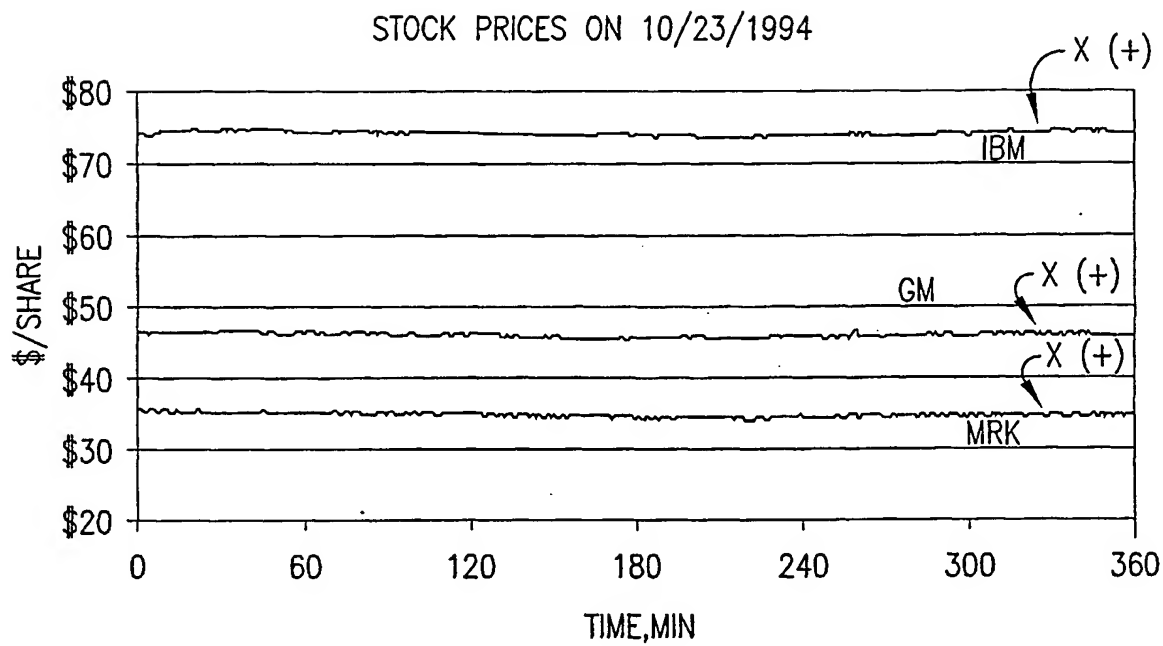
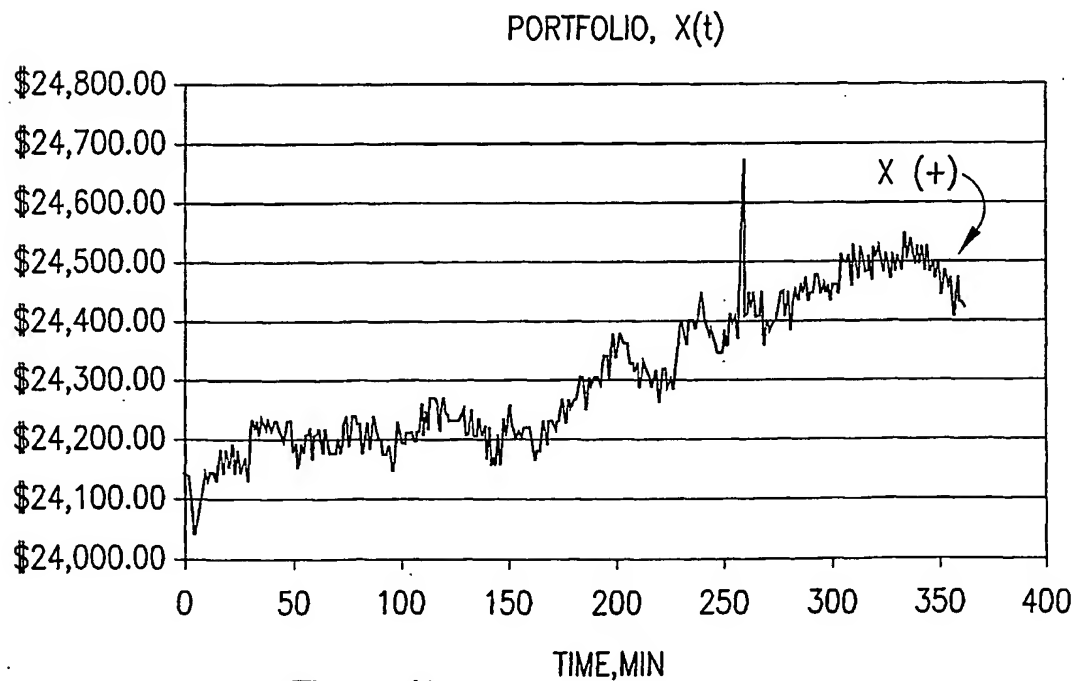


Fig. 3.

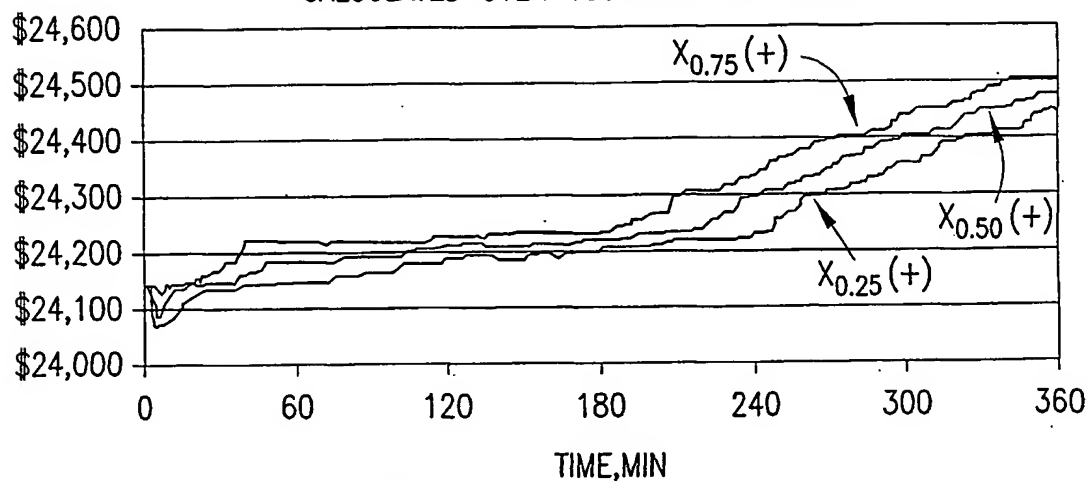
3/11

*Fig. 4.**Fig. 5.*

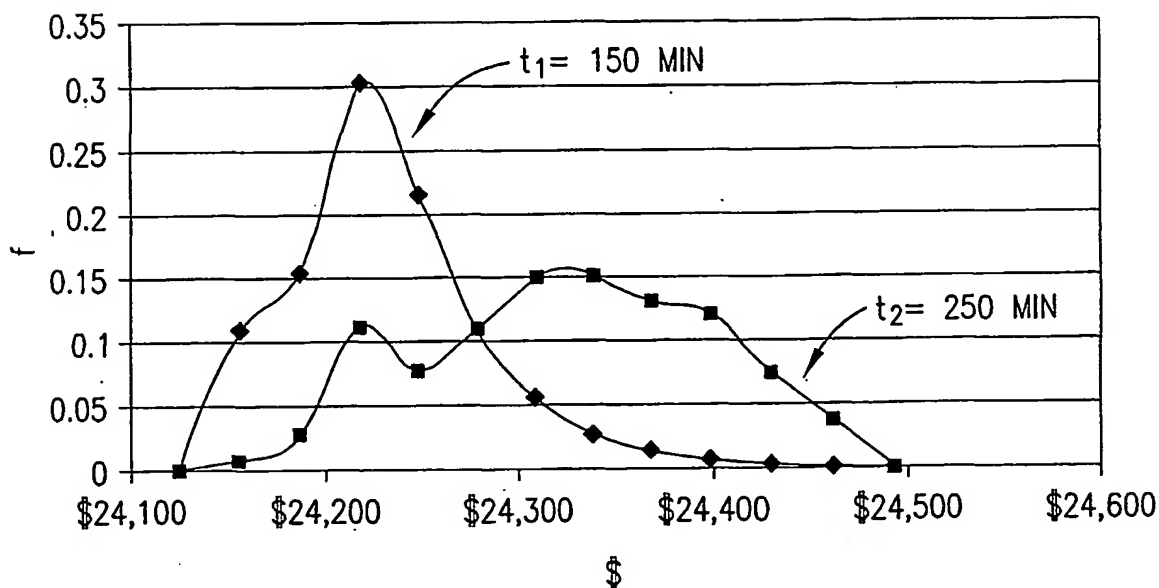


4/11

$X_{0.25}(t)$ ,  $X_{0.50}(t)$  AND  $X_{0.75}(t)$   
CALCULATED OVER 100-MIN TIME-SCALE

*Fig. 6.*

TRUE DENSITIES CALCULATED FOR THE PAST 100 MIN  
- $f(t_1, w)$  AND  $f(t_2, w)$

*Fig. 7.*

5/11

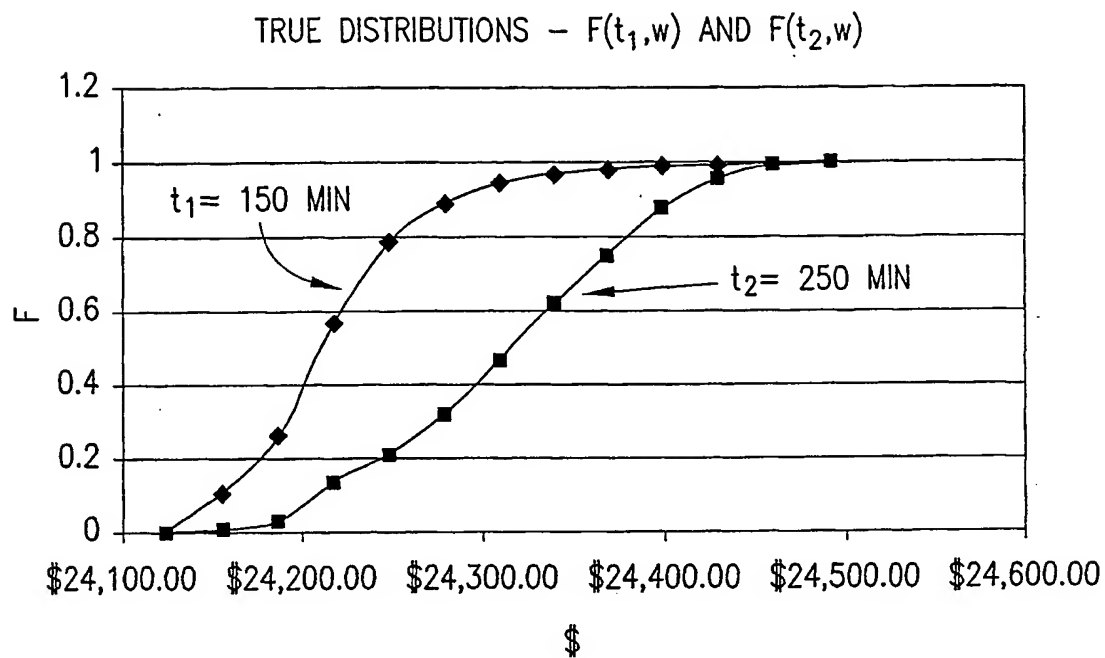


Fig. 8.

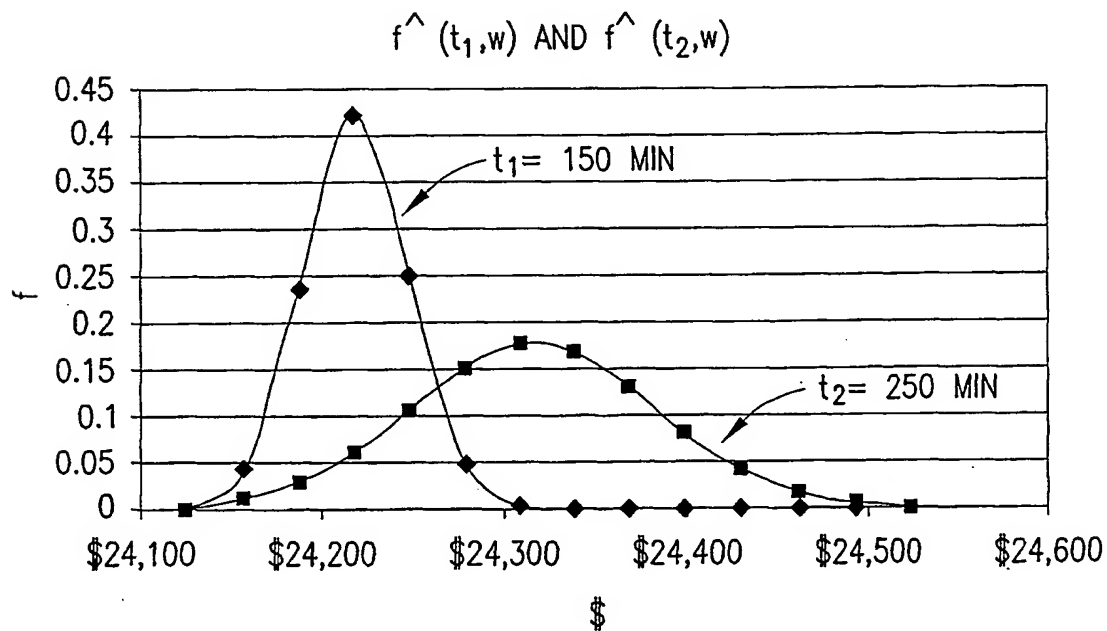


Fig. 9.

6/11

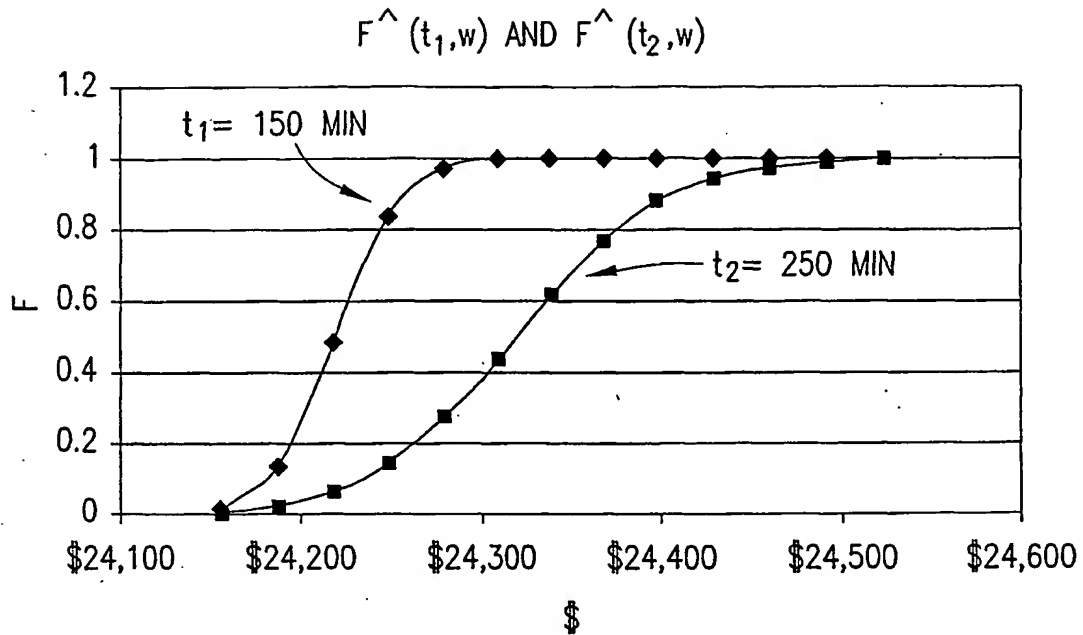


Fig. 10.

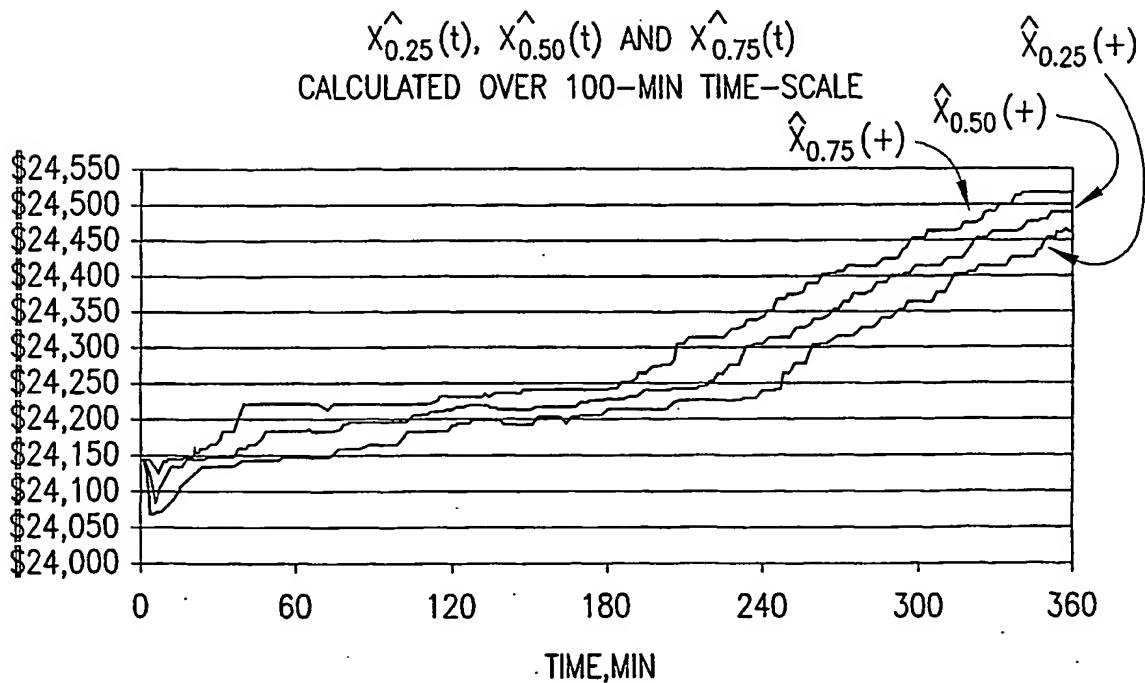


Fig. 11.

7/11

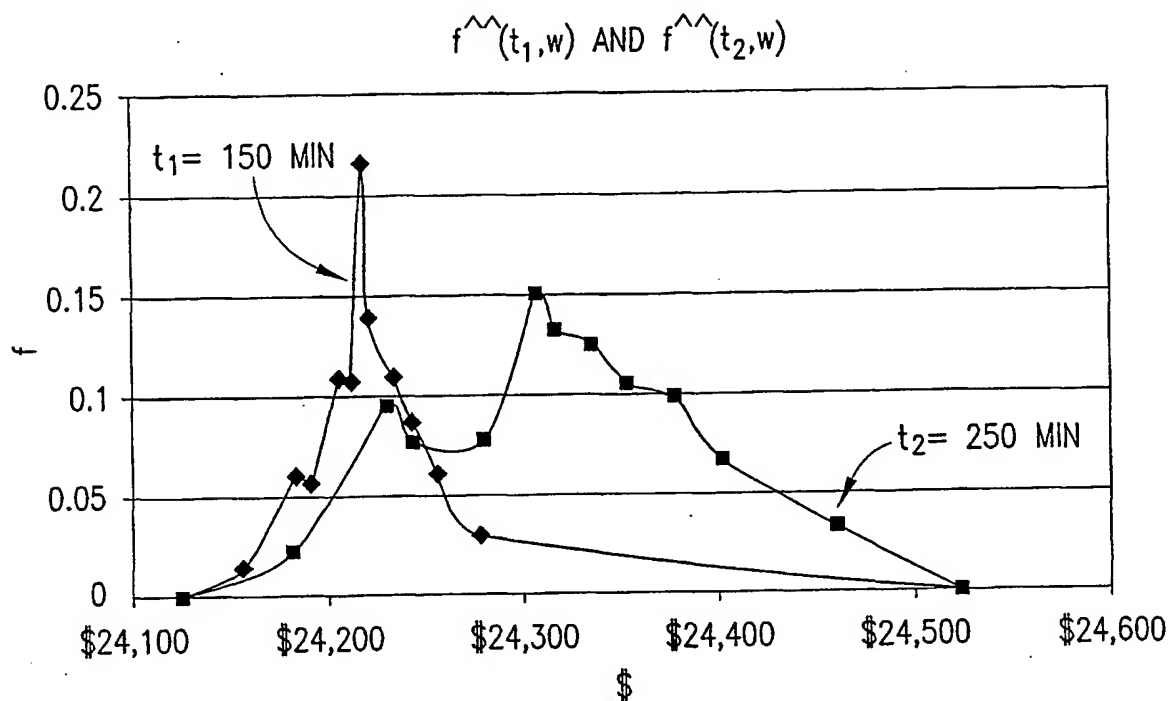


Fig. 12.

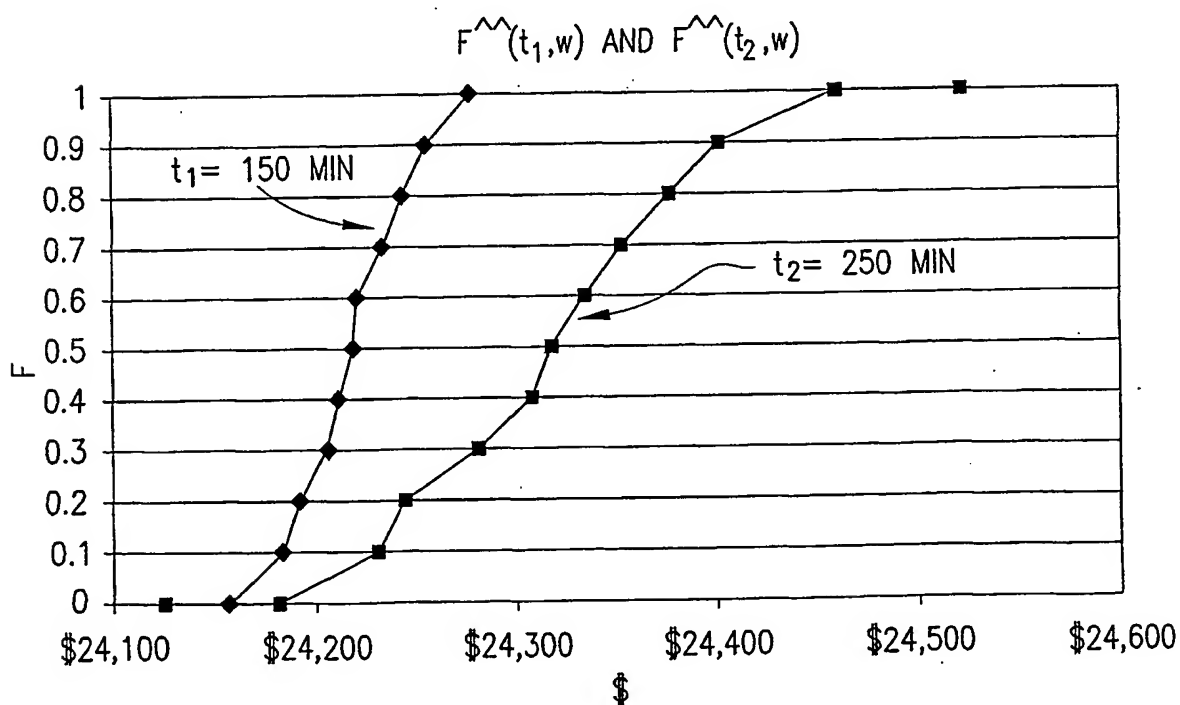


Fig. 13.

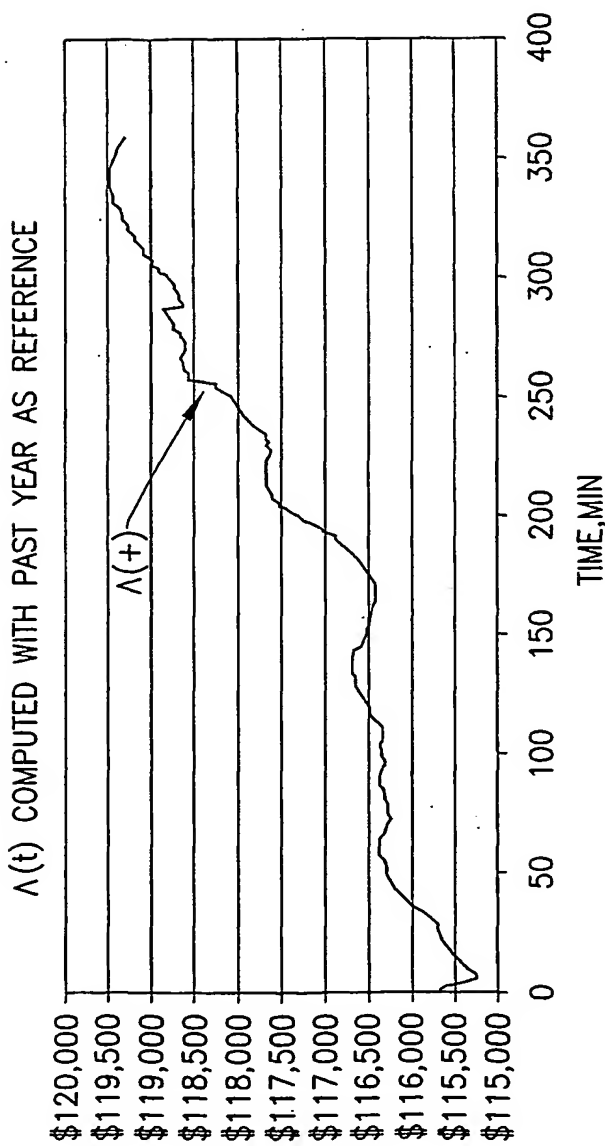


Fig. 14.

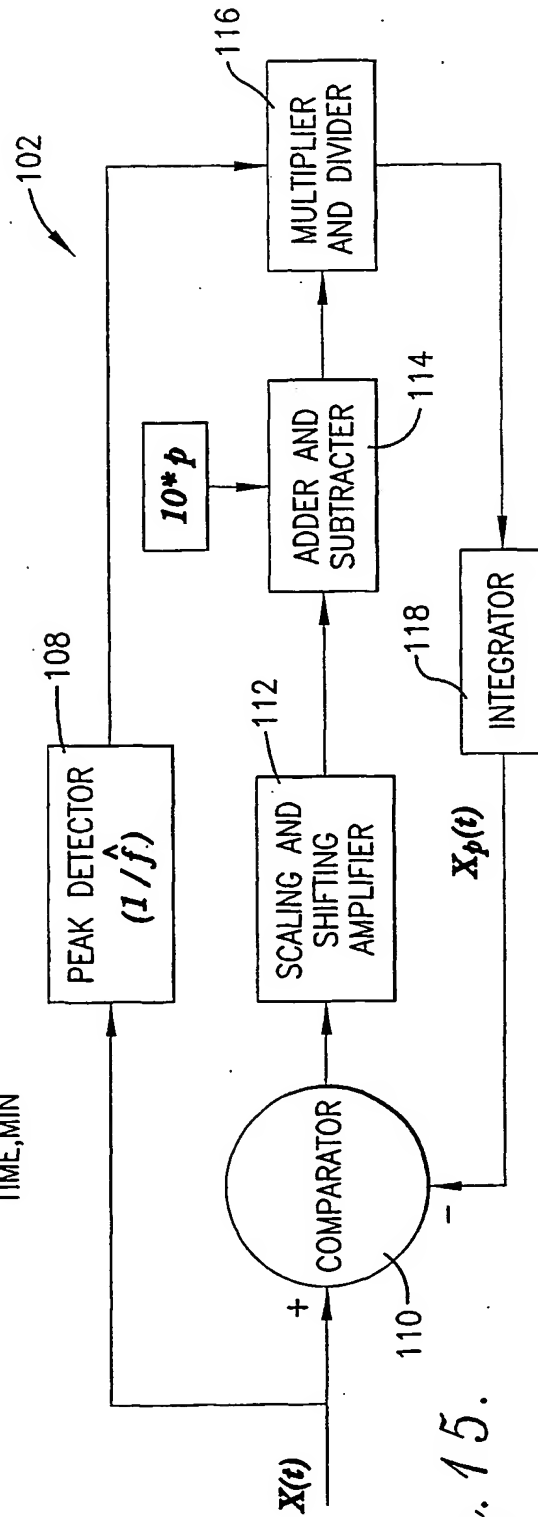


Fig. 15.

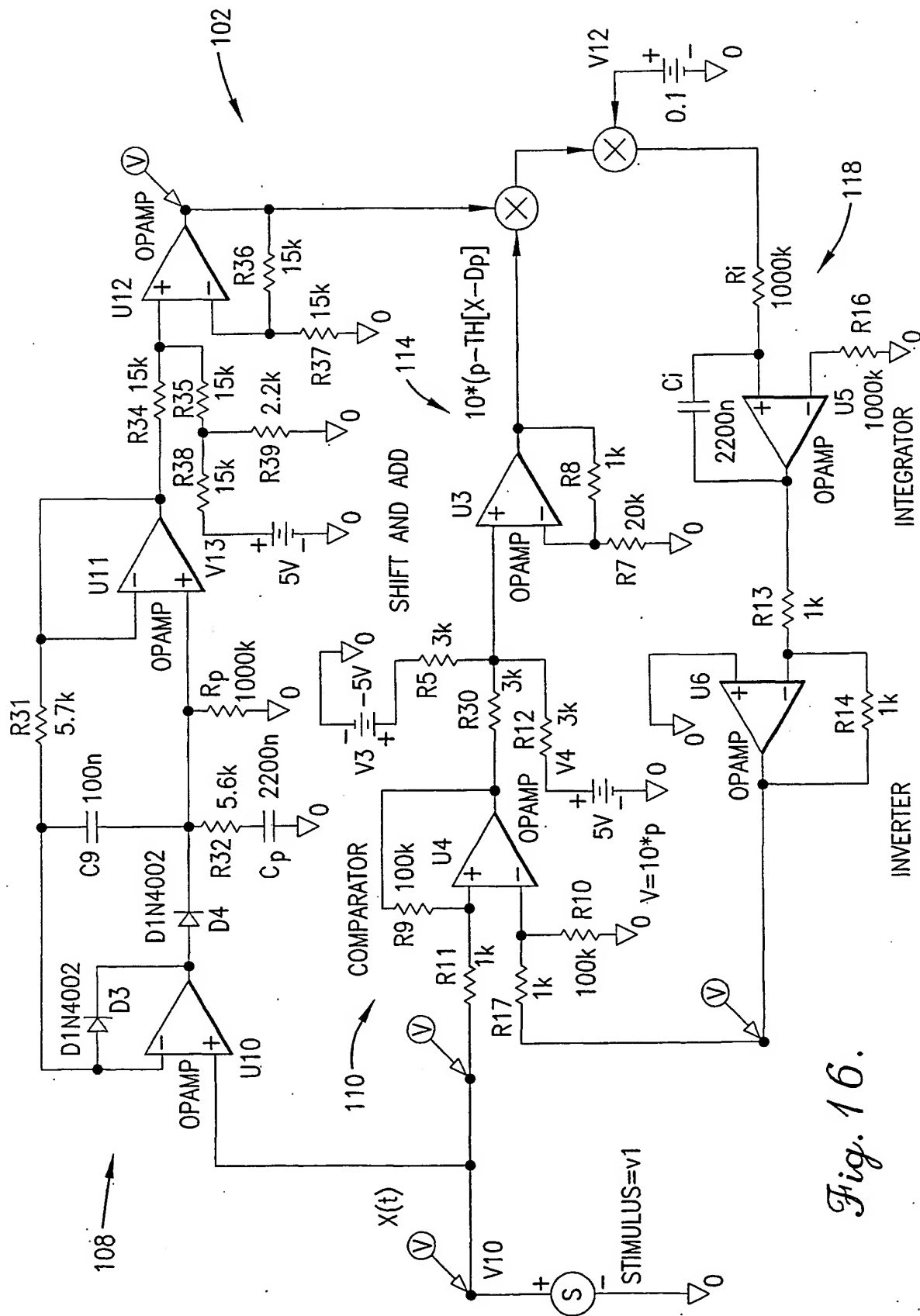
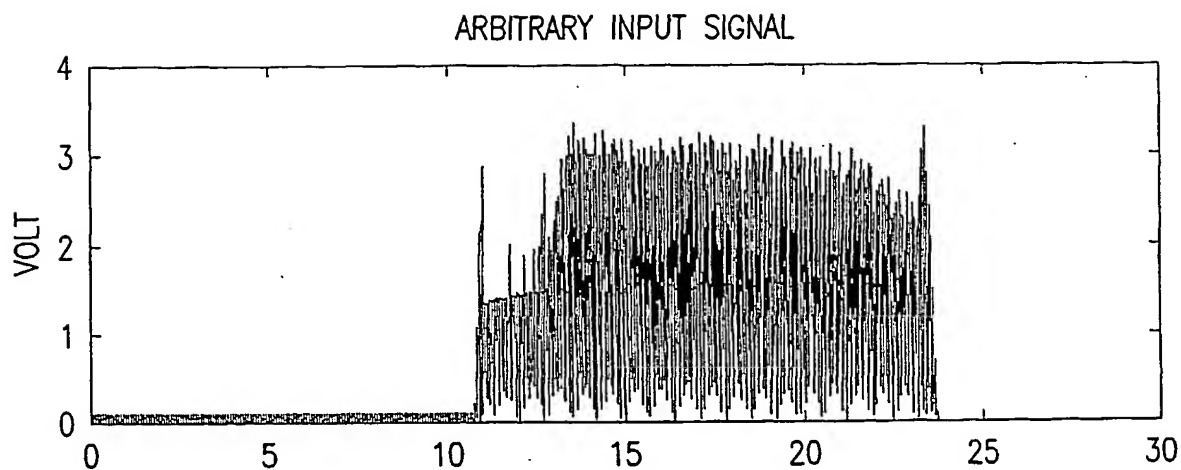
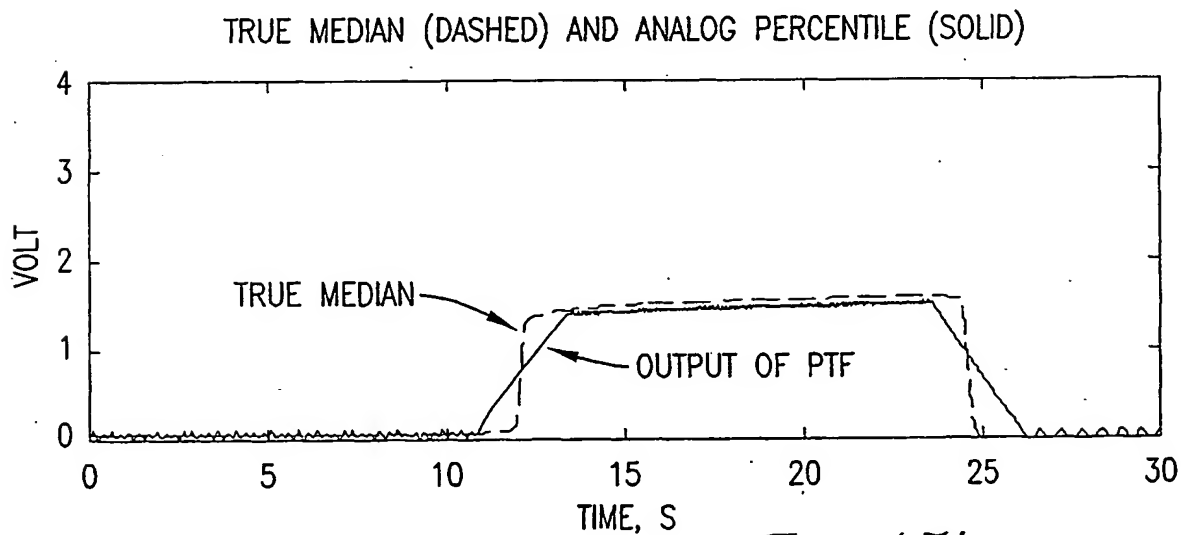
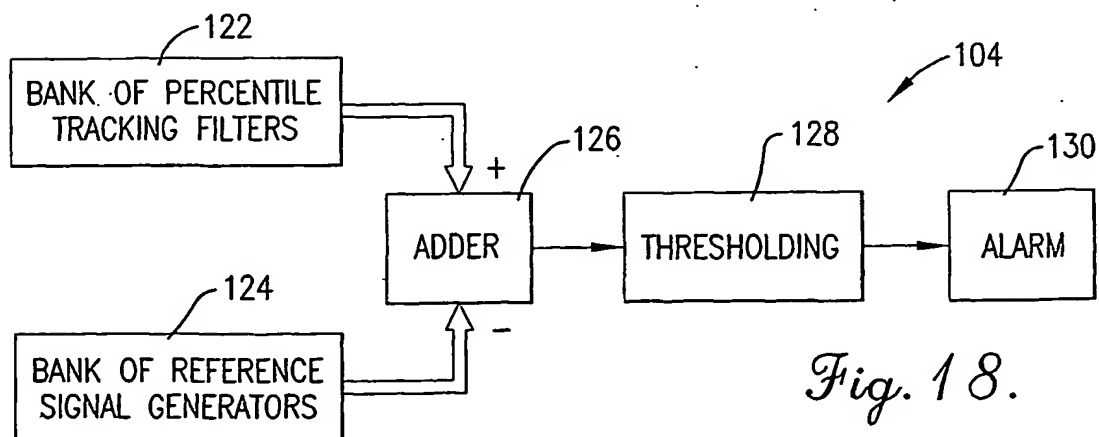
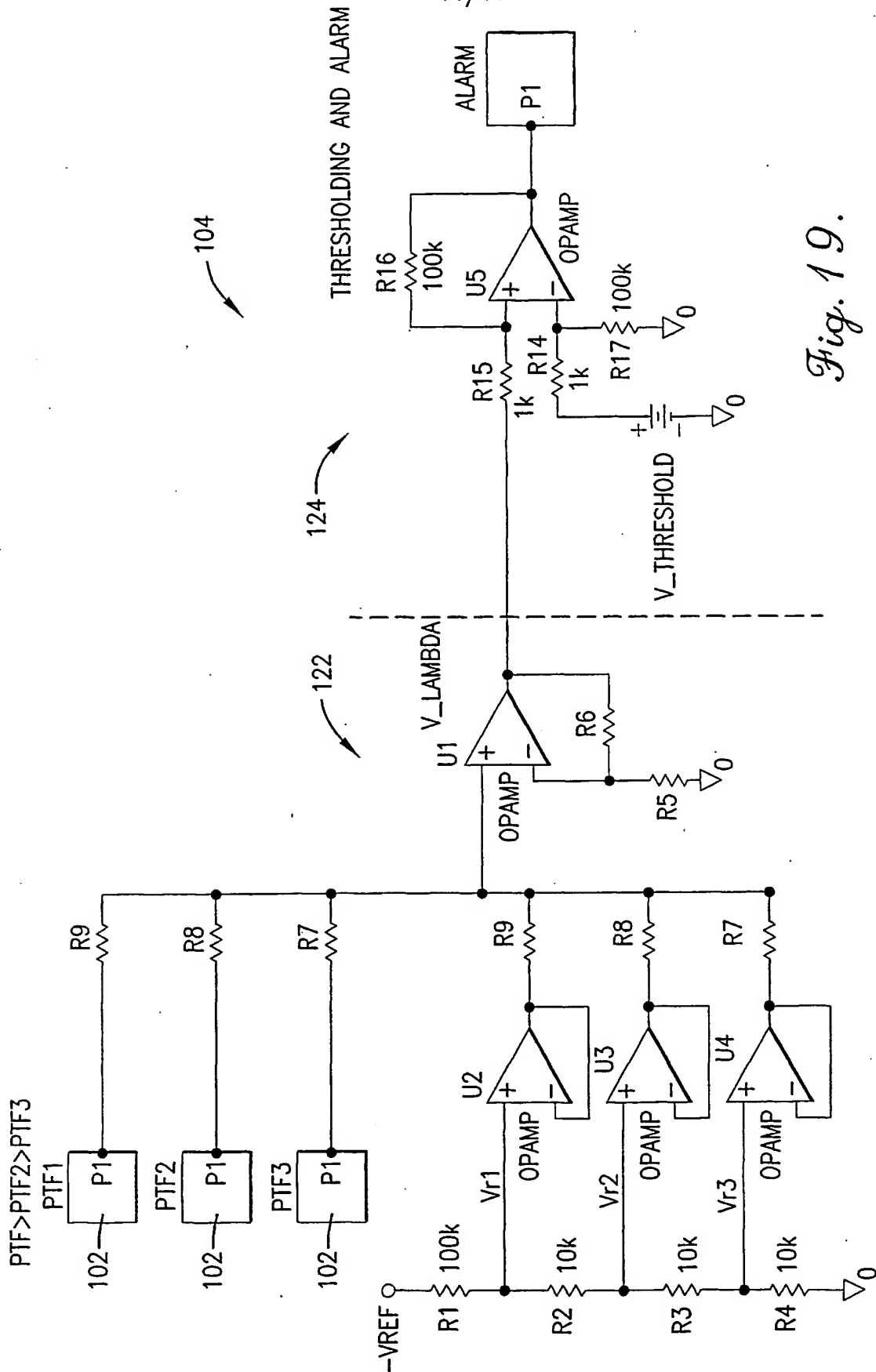


Fig. 16.

10/11

*Fig. 17a.**Fig. 17b.**Fig. 18.*

11/11





## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US01/10677

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(7) : G06F 17/13, 17/14

US CL : 702/66, 124, 189, 194; 324/76.12, 76.22, 77.11

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 702/66, 124, 189, 194; 324/76.12, 76.22, 77.11

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
NONEElectronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
APS search terms: signal, predict, change, percent, tracking, filter, estimate, time, compare, output, background**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A, P	US 6,081,144 A (USUKI et al.) 27 June 2000 (27.06.2000), column 5, lines 4-6.	1-17
A	US 5,583,963 A (LOZACH) 10 December 1996 (10.12.1996), column 3, lines 30-37.	1-17
A	US 5,519,605 A (CAWFIELD) 21 May 1996 (21.05.1996), column 10, lines 11-19.	1-17
A	US 5,347,446 A (IINO et al.) 13 September 1994 (13.09.1994), column 6, lines 37-51.	1-17
A	US 4,998,051 A (ITO) 05 March 1991 (05.03.1991), column 1, line 55 through column 2, line 36.	1-17
A	US 4,663,703 A (AXELBY et al.) 05 May 1987 (05.05.1987), Figure 2.	1-17
A	US 4,791,548 A (YOSHIKAWA et al.) 13 December 1988 (13.12.1988), column 6, lines 41-45.	1-17



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier application or patent published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

03 June 2001 (03.06.2001)

Date of mailing of the international search report

30 JUL 2001

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703)305-3230

Authorized officer

Marc Hoff

Telephone No. 703-308-1782

**This Page Blank (uspto)**